Gli Universali Linguistici di Greenberg: Un'Analisi Approfondita delle Implicazioni tra Strutture e Pattern Cross-Linguistici dell'Ordine delle Parole

Introduzione Storico-Metodologica: Joseph Greenberg e la Rivoluzione della Linguistica Tipologica

La linguistica del ventesimo secolo è stata profondamente e irrevocabilmente segnata dal contributo intellettuale di Joseph Harold Greenberg (1915-2001). La sua opera non ha semplicemente arricchito il panorama degli studi linguistici, ma ha radicalmente trasformato l'intero paradigma epistemologico attraverso cui concepiamo la relazione dialettica tra la diversità superficiale e l'unità profonda del linguaggio umano. Prima del suo intervento metodologicamente rivoluzionario, la tipologia linguistica era largamente dominata da approcci descrittivi e tassonomici che classificavano le lingue in categorie discrete sulla base di criteri morfologici relativamente superficiali, distinguendo tra lingue isolanti, agglutinanti, flessive e polisintetiche, senza però indagare sistematicamente le correlazioni strutturali tra i diversi fenomeni grammaticali e senza interrogarsi sulle ragioni profonde che potevano spiegare la distribuzione non casuale delle proprietà linguistiche.

Greenberg ha introdotto un metodo empirico rigoroso e quantitativamente orientato, basato sull'analisi comparativa di ampie campionature di lingue provenienti da famiglie genetiche diverse e da aree geografiche distanti, permettendo di identificare pattern ricorrenti che trascendono le contingenze storiche specifiche e le influenze areali derivanti dal contatto linguistico. La sua metodologia rappresentava un punto di equilibrio sofisticato tra l'empirismo radicale, che privilegia l'osservazione dei dati concreti, e la ricerca di generalizzazioni teoricamente significative che possano illuminare i principi organizzativi sottostanti alla facoltà linguistica umana.

Il suo articolo pioneristico del 1963, intitolato "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements", pubblicato nel volume collettaneo "Universals of Language" da lui stesso curato, rappresenta indiscutibilmente una pietra miliare nella storia della linguistica teorica e tipologica. In questo lavoro seminale, Greenberg ha presentato quarantacinque universali linguistici basati sull'analisi di un campione di trenta lingue accuratamente selezionate per rappresentare la massima diversità tipologica possibile all'interno dei vincoli pratici di una ricerca condotta da un singolo studioso. Questo campione includeva lingue provenienti dalle famiglie indoeuropea (come il serbo-croato, il greco, l'italiano, il norvegese), semitica (come l'ebraico e il berbero), uralica (come il finlandese), altaica (come il turco e il mongolo), sino-tibetana (come il birmano e il cinese mandarino), Niger-Congo (come lo yoruba e lo swahili), austronesiana (come il malese e il giavanese), e molte altre famiglie meno conosciute ma cruciali per garantire la diversità del

campione, come le lingue native americane (quechua, zapoteco), le lingue australiane, e le lingue della Nuova Guinea.

La metodologia innovativa di Greenberg consisteva nell'osservare non solo quali strutture grammaticali esistessero nelle lingue del mondo in termini assoluti, ma soprattutto quali combinazioni di strutture tendessero a co-occorrere sistematicamente e quali invece risultassero estremamente rare o del tutto assenti nella tipologia cross-linguistica. Questo approccio si basava sul riconoscimento che la distribuzione delle proprietà grammaticali non è casuale ma mostra pattern statistici significativi che richiedono una spiegazione teorica. La presenza di tali pattern suggerisce l'esistenza di vincoli sottostanti, che possono derivare da proprietà innate della facoltà linguistica umana, da principi cognitivi generali del processamento dell'informazione, da requisiti comunicativi universali, o da tendenze sistematiche nel cambiamento linguistico diacronico.

Il concetto di universale linguistico, nella formulazione metodologicamente sofisticata di Greenberg, si allontana decisivamente dalle concezioni assolute e deterministiche che caratterizzavano una parte significativa della tradizione precedente, in particolare quella derivante dalla linguistica razionalista del diciassettesimo e diciottesimo secolo. Greenberg distingue accuratamente tra universali assoluti, che si applicano senza eccezioni a tutte le lingue conosciute e che definiscono vincoli invalicabili sullo spazio delle grammatiche possibili, e universali statistici o tendenze universali, che descrivono pattern altamente frequenti e statisticamente significativi ma che ammettono eccezioni limitate e motivate. Questa distinzione metodologica si rivela cruciale perché consente di catturare generalizzazioni empiricamente robuste e teoricamente significative anche in presenza di controesempi isolati, evitando così il rischio di scartare pattern genuini semplicemente perché non soddisfano un criterio di universalità assoluta che potrebbe essere empiricamente troppo restrittivo.

La maggior parte degli universali identificati da Greenberg appartiene alla seconda categoria, quella degli universali statistici o tendenziali, riflettendo la consapevolezza metodologica che il linguaggio umano, pur mostrando notevoli regolarità che richiedono una spiegazione teorica, conserva sempre margini di variazione e complessità che sfuggono a schematizzazioni troppo rigide e deterministiche. Questa prospettiva probabilistica e graduale rappresenta un'anticipazione notevole di approcci più recenti nella linguistica teorica che enfatizzano il carattere stocastico e variabile dei fenomeni linguistici.

La Natura Logico-Formale degli Universali Implicazionali: La Struttura "Se X Allora Y" e le Sue Implicazioni Teoriche

Il contributo forse più originale, influente e teoricamente profondo di Greenberg risiede nell'identificazione sistematica e nella formalizzazione degli universali implicazionali, ovvero quegli universali che stabiliscono relazioni di dipendenza logica asimmetrica tra diverse proprietà grammaticali. La forma canonica di questi universali può essere espressa formalmente come un'implicazione materiale nella logica proposizionale: $X \rightarrow Y$, che si legge "se una lingua possiede la proprietà X, allora possiede anche la proprietà Y", oppure,

equivalentemente nella semantica dell'implicazione, "la presenza di X implica la presenza di Y". In termini di teoria degli insiemi, questa relazione può essere rappresentata come $L(X) \subseteq L(Y)$, dove $L(X)$ indica l'insieme delle lingue che possiedono la proprietà X e $L(Y)$ l'insieme delle lingue che possiedono la proprietà Y. L'implicazione afferma che l'insieme delle lingue con X è un sottoinsieme (proprio o improprio) dell'insieme delle lingue con Y.

Questa struttura logica è radicalmente diversa da quella degli universali assoluti semplici, che affermano semplicemente "tutte le lingue possiedono la proprietà Z", formalizzabile come $\forall l \in L: Z(l)$, dove L rappresenta l'insieme di tutte le lingue naturali umane. Gli universali implicazionali non negano necessariamente l'esistenza di lingue che possiedono Y senza possedere X, cioè ammettono $L(Y) \setminus L(X) \neq \varnothing$, ma affermano che la combinazione "X senza Y" è impossibile o estremamente rara, ovvero $L(X) \setminus L(Y) = \varnothing$ o $|L(X) \setminus L(Y)| \ll |L(X)|$, dove i simboli di barra inversa denotano la differenza insiemistica e le barre verticali denotano la cardinalità degli insiemi.

La potenza euristica e teorica di questo tipo di universale deriva fondamentalmente dal fatto che esso non si limita a descrivere proprietà isolate e atomiche delle lingue, considerate indipendentemente l'una dall'altra, ma cattura relazioni sistemiche e strutturalmente significative tra diversi aspetti della grammatica che potrebbero apparire superficialmente non correlati. Un universale implicazionale ci dice che certi fenomeni linguistici non sono variabili indipendenti nello spazio tipologico ma sono legati da connessioni profonde che possono riflettere principi cognitivi fondamentali nella strutturazione e nell'elaborazione dell'informazione linguistica, vincoli comunicativi universali derivanti dalla funzione del linguaggio come sistema di trasmissione dell'informazione, requisiti di processamento che facilitano la produzione e la comprensione in tempo reale, o anche semplici conseguenze logiche della struttura grammaticale interna.

Quando osserviamo empiricamente che la presenza di una certa struttura X tende sistematicamente a co-occorrere con un'altra struttura Y attraverso lingue di famiglie genetiche diverse e di aree geografiche distanti, siamo naturalmente portati a chiederci quali fattori causali o motivazionali causino o favoriscano questa correlazione non casuale. Le risposte teoriche possibili a questa domanda fondamentale includono: primo, vincoli cognitivi innati sulla strutturazione dell'informazione linguistica, derivanti dalla natura specifica della facoltà linguistica umana come sistema computazionale biologicamente determinato; secondo, principi di economia cognitiva generale nel processamento linguistico, che favoriscono configurazioni grammaticali che minimizzano il carico computazionale durante la produzione e la comprensione; terzo, esigenze comunicative universali legate alla necessità di trasmettere informazione in modo efficiente e non ambiguo; quarto, tendenze sistematiche nei processi di cambiamento linguistico diacronico, che favoriscono certi percorsi evolutivi rispetto ad altri; quinto, semplici conseguenze logiche della struttura grammaticale, dove la presenza di X rende necessaria o altamente probabile la presenza di Y per ragioni di coerenza interna del sistema.

Per comprendere più concretamente e in maggiore profondità la logica strutturale degli universali implicazionali, consideriamo un esempio paradigmatico, l'Universale 2 di Greenberg, che ha ricevuto notevole attenzione nella letteratura tipologica successiva: "In lingue con preposizioni, il genitivo segue quasi sempre il nome governante, mentre in lingue con postposizioni il genitivo precede quasi sempre il nome governante". Questo universale

stabilisce una correlazione sistematica e robusta tra due fenomeni apparentemente distinti e operanti in domini sintattici diversi: da un lato, la posizione relativa delle adposizioni (elementi grammaticali che marcano le relazioni grammaticali e semantiche dei nomi, come le preposizioni in italiano "di", "a", "da", "in", "con", "su", "per", "tra", "fra") rispetto ai loro complementi nominali; dall'altro lato, la posizione relativa dei modificatori genitivi (che esprimono possesso, relazioni part-whole, o altre relazioni associative) rispetto ai nomi che modificano.

La forma implicazionale può essere espressa più precisamente attraverso due implicazioni correlate che costituiscono le due parti dell'universale: prima implicazione, se una lingua ha preposizioni (cioè adposizioni che precedono sistematicamente il loro complemento nominale), allora il genitivo seguirà il nome che modifica, generando costruzioni come "il libro di Marco" in italiano, "the book of John" in inglese, "le livre de Jean" in francese, dove il modificatore genitivo appare in posizione postnominale; seconda implicazione, se invece una lingua ha postposizioni (cioè adposizioni che seguono sistematicamente il loro complemento nominale), allora il genitivo precederà il nome che modifica, generando costruzioni come "Mehmet'in kitabı" in turco (letteralmente "Mehmet-GENITIVO libro-POSSESSIVO-3SG"), "Tarō no hon" in giapponese (letteralmente "Taro GENITIVO libro"), dove il modificatore genitivo appare in posizione prenominale.

Questa correlazione non è ovvia a priori e richiede una spiegazione teorica articolata e empiricamente motivata. Perché la posizione delle adposizioni rispetto ai loro complementi dovrebbe influenzare sistematicamente la posizione dei modificatori genitivi rispetto ai nomi che modificano? Quale principio organizzativo sottostante può spiegare questa apparente "simpatia" strutturale tra domini grammaticali distinti? Una possibile spiegazione, che Greenberg stesso suggerisce in forma embrionale e che è stata ampiamente sviluppata, raffinata e formalizzata nella letteratura tipologica e teoretica successiva, si basa sul concetto di armonia direzionale o consistenza nell'ordine dei costituenti, talvolta denominato anche principio di branching consistente o principio della direzionalità della testa.

L'idea fondamentale alla base di questo principio è che le lingue tendono a essere coerenti e sistematiche nella direzione in cui le teste sintattiche (elementi governanti che determinano le proprietà categoriali della frase) si collocano rispetto ai loro dipendenti o complementi (elementi governati che completano semanticamente la testa). Se una lingua manifesta una preferenza per l'ordine "head-initial" o testa iniziale, collocando sistematicamente le teste sintattiche prima dei loro dipendenti attraverso diverse proiezioni categoriali, questa tendenza armonica si manifesterà in modo consistente sia nel dominio delle adposizioni (dove la preposizione, che funziona come la testa sintattica della frase adposizionale, precede il suo complemento nominale) sia nel dominio della modificazione nominale (dove il nome, che funziona come la testa sintattica del sintagma nominale esteso, precede il suo modificatore genitivo). Viceversa, le lingue che manifestano una preferenza per l'ordine "head-final" o testa finale mostreranno la tendenza opposta e speculare in entrambi i domini sintattici.

La forza predittiva empirica degli universali implicazionali emerge con particolare chiarezza quando consideriamo sistematicamente tutte le quattro combinazioni logicamente possibili tra le due proprietà binarie in questione, organizzate in una tabella di contingenza due per due. Le quattro combinazioni possibili sono: primo, preposizioni combinate con genitivo

postnominale (come manifestato nelle lingue romanze italiane, francese, spagnolo, portoghese, rumeno, nelle lingue germaniche come inglese, tedesco standard, olandese, svedese, e in molte altre lingue indoeuropee e non indoeuropee con ordine VO); secondo, postposizioni combinate con genitivo prenominale (come manifestato nelle lingue altaiche turco, mongolo, tunguso, nelle lingue uralico-finniche finlandese, estone, ungherese, nelle lingue giapponese, coreano, basco, e in molte altre lingue con ordine OV); terzo, la combinazione teoricamente possibile ma empiricamente rara di preposizioni combinate con genitivo prenominale; quarto, la combinazione teoricamente possibile ma empiricamente rara di postposizioni combinate con genitivo postnominale.

L'Universale 2 prevede con precisione che le combinazioni terza e quarta siano estremamente rare o praticamente inesistenti nella distribuzione tipologica mondiale, e l'evidenza empirica accumulata attraverso decenni di ricerca tipologica su campioni progressivamente più ampi conferma ampiamente questa previsione con notevole robustezza statistica. Quando si esaminano sistematicamente le lingue del mondo utilizzando database tipologici come il World Atlas of Language Structures (WALS), che contiene informazioni su oltre 2600 lingue, si scopre che la stragrande maggioranza delle lingue rientra nei tipi primo o secondo, mentre i tipi terzo e quarto sono talmente rari che possono essere considerati marginali o virtualmente inesistenti, a seconda di come si interpretano i casi ambigui e le lingue che mostrano variazione interna o ordini multipli.

Possiamo quantificare la forza di questa correlazione utilizzando misure statistiche appropriate. Se definiamo P come la probabilità che una lingua abbia preposizioni, Po come la probabilità che una lingua abbia postposizioni (con P + Po ≈ 1,0 se trascuriamo le lingue senza adposizioni), G_post come la probabilità che il genitivo sia postnominale, e G_pre come la probabilità che il genitivo sia prenominale (con G_post + G_pre ≈ 1,0 se trascuriamo casi ambigui), allora l'Universale 2 afferma che P(G_post | P) ≈ 0,85-0,95 e P(G_pre | Po) ≈ 0,85-0,95, dove P(A | B) denota la probabilità condizionale di A dato B. In altre parole, data una lingua con preposizioni, la probabilità che abbia genitivo postnominale è molto alta (tra 0,85 e 0,95 secondo diverse stime basate su diversi campioni), e data una lingua con postposizioni, la probabilità che abbia genitivo prenominale è similmente molto alta.

Un aspetto logicamente cruciale degli universali implicazionali è la loro asimmetria intrinseca e irriducibile. Se affermiamo formalmente X → Y, non stiamo simultaneamente affermando Y → X. In termini di logica proposizionale, l'implicazione è unidirezionale e non equivale alla biimplicazione o doppia implicazione X ↔ Y. Questo significa che mentre la presenza di X garantisce (o rende altamente probabile nel caso di universali statistici) la presenza di Y, la presenza di Y non garantisce automaticamente la presenza di X. Nel caso specifico dell'Universale 2, se osserviamo che una lingua ha preposizioni possiamo prevedere con alta confidenza statistica (dell'ordine del 85-95%) che avrà il genitivo postnominale, ma se osserviamo che una lingua ha il genitivo postnominale non possiamo automaticamente concludere con la stessa confidenza che avrà preposizioni, perché potrebbe essere una lingua che manca del tutto di adposizioni come categoria grammaticale distinta, o che manifesta proprietà miste in questa dimensione.

Questa asimmetria logica riflette il fatto profondo che gli universali implicazionali identificano e codificano relazioni di dipendenza asimmetriche tra proprietà grammaticali, dove una proprietà (l'antecedente dell'implicazione) è in qualche senso più "marcata" o più specifica

dell'altra (il conseguente dell'implicazione). La proprietà X può essere vista come implicante vincoli o requisiti strutturali più stringenti che automaticamente garantiscono la presenza di Y, mentre Y può esistere indipendentemente in assenza di X.

Gli universali implicazionali possono anche essere concatenati e composti in catene implicazionali più lunghe, creando gerarchie implicazionali complesse che rivelano strutture di dipendenza multidimensionali. Se abbiamo tre proprietà A, B e C tali che A implica B (formalizzato come A → B) e B implica C (formalizzato come B → C), otteniamo per transitività dell'implicazione una gerarchia o scala implicazionale A → B → C, talvolta notata compattamente come A > B > C, dove il simbolo > rappresenta la relazione di implicazione direzionale. Questo tipo di gerarchia stabilisce una scala di marcatezza tipologica o una gerarchia di accessibilità: le lingue che possiedono la proprietà più marcata o più restrittiva (A) possiederanno necessariamente anche quelle meno marcate o meno restrittive (B e C), mentre è possibile trovare lingue che possiedono solo C, o che possiedono C e B senza possedere A.

Queste gerarchie implicazionali si rivelano particolarmente utili e teoreticamente illuminanti per diversi scopi: primo, per comprendere la distribuzione sistematica delle proprietà grammaticali attraverso le lingue del mondo e identificare quali combinazioni sono attestate versus quali sono lacune tipologiche; secondo, per formulare ipotesi testabili sull'acquisizione linguistica da parte dei bambini, con la previsione che le proprietà meno marcate (più in basso nella gerarchia) saranno acquisite prima o più facilmente di quelle più marcate; terzo, per modellare il cambiamento linguistico diacronico, con l'ipotesi che i percorsi di cambiamento tendano a seguire la gerarchia, con perdite che procedono dall'alto verso il basso (perdita di proprietà marcate prima di quelle non marcate) e innovazioni che procedono dal basso verso l'alto.

L'Ordine delle Parole e i Pattern Tipologici Fondamentali: Un'Analisi Quantitativa della Distribuzione Cross-Linguistica

L'ordine lineare dei costituenti maggiori costituisce indiscutibilmente uno degli aspetti più centrali, visibili e immediatamente osservabili della variazione tipologica tra le lingue del mondo. Mentre alcune proprietà grammaticali di natura più astratta richiedono analisi sintattiche sofisticate e teoricamente informate per essere identificate e caratterizzate precisamente, l'ordine sequenziale in cui vengono disposti i costituenti maggiori di una frase transitiva semplice è immediatamente osservabile attraverso l'ispezione superficiale dei dati e facilmente quantificabile mediante procedure di conteggio dirette. Non sorprende quindi che gli universali formulati da Greenberg dedicassero un'attenzione particolare e privilegiata a questo dominio empirico, e che molti degli universali più noti, influenti e ampiamente discussi nella letteratura tipologica riguardassero proprio le correlazioni sistematiche tra diversi aspetti dell'ordine dei costituenti.

Il punto di partenza logico e metodologico dell'analisi di Greenberg è l'ordine relativo dei tre costituenti fondamentali della frase transitiva semplice, che rappresentano gli elementi minimali necessari per descrivere un evento transitivo prototipico: il soggetto grammaticale

(denotato convenzionalmente con la lettera S), il verbo o predicato verbale (denotato con V), e l'oggetto diretto o paziente (denotato con O). Dal punto di vista puramente logico-combinatorio, esistono esattamente sei permutazioni possibili per questi tre elementi distinti: SOV (soggetto-oggetto-verbo), SVO (soggetto-verbo-oggetto), VSO (verbo-soggetto-oggetto), VOS (verbo-oggetto-soggetto), OVS (oggetto-verbo-soggetto) e OSV (oggetto-soggetto-verbo).

Se gli ordini fossero distribuiti casualmente e uniformemente, ci aspetteremmo che ciascuno dei sei ordini fosse approssimativamente equiprobabile, con una probabilità teorica di 1/6 ≈ 0,167 o circa 16,7%. Tuttavia, l'analisi empirica condotta originariamente da Greenberg sul suo campione di trenta lingue, e ampiamente confermata, raffinata e quantificata con maggiore precisione da ricerche tipologiche successive basate su campioni molto più ampi e statisticamente più rappresentativi, mostra in modo inequivocabile che questi sei tipi non sono affatto equamente distribuiti nelle lingue del mondo ma manifestano una distribuzione altamente asimmetrica e statisticamente significativa.

Tre ordini emergono come nettamente dominanti nella tipologia mondiale: SOV (soggetto-oggetto-verbo), caratteristico delle lingue altaiche, dravidiche, uralico-finniche, giapponese, coreano, basco, e molte altre lingue sparse geograficamente; SVO (soggetto-verbo-oggetto), caratteristico delle lingue romanze, germaniche, slave, e molte altre lingue indoeuropee, oltre a numerose lingue non indoeuropee sparse nel mondo; e VSO (verbo-soggetto-oggetto), meno frequente dei primi due ma comunque ben attestato nelle lingue celtiche, semitiche, e in varie lingue austronesiane e native americane.

Le stime quantitative più recenti e metodologicamente rigorose, basate su database tipologici molto più ampi di quello disponibile a Greenberg negli anni 1960, in particolare il World Atlas of Language Structures che contiene dati sistematici su oltre 1500 lingue per questa variabile specifica, suggeriscono la seguente distribuzione approssimativa: circa il 45% delle lingue campionate è di tipo SOV (con stime che variano tra 41% e 48% a seconda del campione specifico e dei criteri di classificazione), circa il 42% è di tipo SVO (con stime tra 38% e 45%), e circa il 9% è di tipo VSO (con stime tra 7% e 12%). Gli altri tre ordini teoricamente possibili (VOS, OVS, OSV) costituiscono insieme meno del 4% delle lingue conosciute e documentate, con distribuzioni approssimative di VOS circa 3%, OVS meno dell'1%, e OSV estremamente raro con meno dello 0,5%, rappresentando forse una dozzina di lingue su diverse migliaia.

Possiamo esprimere queste frequenze in termini di probabilità empiriche stimate da campioni tipologici: $P(SOV) \approx 0,45$, $P(SVO) \approx 0,42$, $P(VSO) \approx 0,09$, $P(VOS) \approx 0,03$, $P(OVS) \approx 0,01$, $P(OSV) < 0,01$. La somma di queste probabilità approssima 1,0 considerando anche una piccola percentuale di lingue con ordine libero o non dominante. Se confrontiamo queste probabilità empiriche con la distribuzione uniforme teorica dove ciascun ordine avrebbe $P = 1/6 \approx 0,167$, possiamo calcolare un test chi-quadrato per valutare la significatività statistica della deviazione dall'uniformità. Il valore chi-quadrato risultante è estremamente alto ($\chi^2 >> 100$ per campioni di dimensione superiore a 1000), con un p-value essenzialmente nullo ($p < 0,0001$), indicando che la distribuzione osservata è altamente significativamente diversa da quella uniforme.

Questa distribuzione altamente asimmetrica e statisticamente significativa solleva immediatamente questioni teoriche profonde e fondamentali che stanno al cuore della linguistica tipologica e della teoria del linguaggio. Perché certi ordini sono così frequenti mentre altri sono così rari? Quali principi cognitivi, comunicativi, strutturali o elaborativi sottostanti favoriscono SOV, SVO e VSO rispetto a VOS, OVS e OSV? Cosa distingue sistematicamente i tre ordini comuni dai tre ordini rari?

Una prima osservazione empirica rilevante e teoricamente significativa è che i tre ordini dominanti (SOV, SVO, VSO) condividono una proprietà strutturale comune che li distingue categoricamente dai tre ordini rari: in tutti e tre gli ordini dominanti, il soggetto grammaticale precede linearmente l'oggetto diretto nella sequenza superficiale (S...O, dove i puntini denotano material interveniente potenzialmente occupato dal verbo). Questo pattern, noto nella letteratura come preferenza universale per l'ordine S prima di O, o più formalmente come la generalizzazione S > O, sembra riflettere principi cognitivi o comunicativi universali profondamente radicati nella natura dell'elaborazione linguistica umana.

Una spiegazione teorica plausibile e ampiamente discussa nella letteratura è che il soggetto, essendo tipicamente e prototipicamente l'agente dell'azione in una frase transitiva (l'entità che inizia e controlla l'evento), e essendo simultaneamente l'entità più topicale o data nel contesto discorsivo precedente (l'entità di cui si sta parlando, che costituisce il punto di partenza informazionale della proposizione), viene naturalmente e preferenzialmente collocato prima dell'oggetto nella serializzazione lineare. L'oggetto, al contrario, rappresenta invece il paziente dell'azione (l'entità che subisce o è affetta dall'evento) e tende ad essere meno topicale e più focale o nuovo dal punto di vista informazionale.

Questa gerarchia strutturale soggetto > oggetto, dove il simbolo > denota precedenza o prominenza, si manifesta non solo nell'ordine lineare superficiale ma anche in molti altri aspetti indipendenti della grammatica universale, fornendo evidenza convergente per la sua validità cognitiva: primo, la codifica morfologica, dove in molte lingue il soggetto riceve marcatura zero mentre l'oggetto è morfologicamente marcato; secondo, l'accordo verbale, dove il verbo concorda primariamente con il soggetto piuttosto che con l'oggetto nella maggior parte delle lingue; terzo, le gerarchie di accessibilità relativa, dove le posizioni sintattiche più alte come il soggetto sono più facilmente relativizzabili di quelle più basse come l'oggetto; quarto, i fenomeni di binding e controllo, dove il soggetto tipicamente ha precedenza nel legare pronomi riflessivi o controllare clausole infinitive.

Un altro aspetto cruciale e teoricamente centrale riguarda la posizione del verbo rispetto ai suoi argomenti nominali, in particolare rispetto all'oggetto diretto. Le lingue si dividono fondamentalmente in due grandi famiglie o macrotipi: primo, quelle in cui il verbo precede il suo oggetto (ordini VO, che includono tipologicamente sia SVO sia VSO), manifestando l'ordine V...O nella sequenza superficiale; secondo, quelle in cui il verbo segue il suo oggetto (ordini OV, dominati numericamente dal tipo SOV ma includendo anche i rari tipi OSV e OVS), manifestando l'ordine O...V.

Questa distinzione binaria tra lingue VO e lingue OV si rivela fondamentale e organizzativa perché correla sistematicamente e robustatamente con numerose altre proprietà indipendenti dell'ordine delle parole in diversi domini grammaticali, dando origine ad alcuni degli universali implicazionali più robusti, meglio documentati e teoricamente influenti

identificati da Greenberg e confermati da tutta la ricerca tipologica successiva. Possiamo concettualizzare il parametro VO/OV come un parametro macrotipologico centrale che organizza gran parte della variazione cross-linguistica nell'ordine dei costituenti.

L'Universale 4, uno dei più celebri, robusti e ampiamente citati nella letteratura tipologica, afferma con notevole precisione predittiva che con una probabilità schiacciante superiore al caso (ovvero con una probabilità significativamente maggiore di 0,5 e tipicamente nell'intervallo 0,85-0,95), le lingue con ordine normale SOV hanno le adposizioni postposte ai nomi che governano, ovvero hanno postposizioni anziché preposizioni come categoria grammaticale dominante o esclusiva. Questo universale stabilisce una correlazione sistematica e robusta tra due domini sintattici apparentemente distinti: primo, l'ordine dei costituenti maggiori della frase a livello della proiezione verbale (dove si posiziona il verbo rispetto all'oggetto nella gerarchia sintagmatica); secondo, l'ordine all'interno del sintagma adposizionale a livello della proiezione adposizionale (dove si posiziona l'adposizione rispetto al suo complemento nominale).

La logica teorica sottostante è ancora una volta quella dell'armonia direzionale o consistenza parametrica: le lingue SOV sono fondamentalmente e primariamente lingue OV, cioè lingue in cui le teste sintattiche (come il verbo V che è la testa della proiezione VP) seguono sistematicamente i loro complementi sintattici (come l'oggetto O che è il complemento del verbo). Questa preferenza strutturale per l'ordine "complemento-testa" o "head-final" si estende coerentemente e armonicamente ad altri domini sintattici oltre quello verbale, incluso in particolare quello delle adposizioni, dove il complemento nominale (il nome N) precede la testa adposizionale (la postposizione P), generando la struttura [NP N] P piuttosto che P [NP N].

Possiamo formalizzare questa correlazione in termini probabilistici: P(Postposizioni | SOV) ≈ 0,90, dove la probabilità condizionale P(A | B) si legge "la probabilità di A dato B". Questa alta probabilità condizionale indica che, dato che una lingua è classificata come SOV, la probabilità che abbia postposizioni è molto alta, intorno al 90% secondo le stime più affidabili basate su campioni ampi. Al contrario, la probabilità che una lingua SOV abbia preposizioni è corrispondentemente bassa: P(Preposizioni | SOV) ≈ 0,10. Questi valori deviano fortemente dalla probabilità di base incondizionata, che sarebbe circa 0,50 se non ci fosse correlazione.

Analogamente e in modo speculare, l'Universale 3 stabilisce che le lingue con ordine VSO dominante hanno quasi sempre preposizioni come tipo di adposizione. Anche qui la spiegazione teorica fa appello al principio di armonia direzionale: le lingue VSO sono fondamentalmente lingue VO, cioè lingue in cui le teste sintattiche precedono sistematicamente i loro complementi. Questa preferenza strutturale per l'ordine "testa-complemento" o "head-initial" si manifesta in modo consistente sia nel dominio verbale (V precede O, generando V [NP O]) sia nel dominio adposizionale (preposizione precede il nome, generando P [NP N]). Formalmente: P(Preposizioni | VSO) ≈ 0,95, indicando una correlazione estremamente robusta.

Gli universali 3 e 4 insieme stabiliscono quindi una correlazione robustissima e bidirezionale tra il parametro fondamentale dell'ordine VO/OV e il tipo di adposizione (preposizioni versus postposizioni), con pochissime eccezioni sistematicamente documentate nelle lingue del

mondo. Possiamo esprimere questa correlazione attraverso una tabella di contingenza teorica, dove le celle diagonali (VO+Prep, OV+Postp) sono densamente popolate mentre le celle off-diagonali (VO+Postp, OV+Prep) sono scarsamente popolate o virtualmente vuote.

Possiamo quantificare la forza di associazione tra queste due variabili binarie usando varie misure statistiche. Il coefficiente phi (φ), una misura di associazione per variabili binarie, può essere calcolato dalla tabella di contingenza 2×2. Per la correlazione VO/OV versus Preposizioni/Postposizioni, stime basate su campioni ampi suggeriscono $\varphi \approx 0{,}80\text{-}0{,}90$, indicando un'associazione molto forte. Il coefficiente phi varia tra -1 (associazione negativa perfetta) e +1 (associazione positiva perfetta), con 0 che indica assenza di associazione. Un valore di 0,80-0,90 indica che la conoscenza del tipo VO/OV predice con alta accuratezza il tipo di adposizione, e viceversa.

L'analisi si complica interessantemente quando consideriamo le lingue SVO, che rappresentano circa il 42% del totale delle lingue del mondo e che costituiscono il secondo tipo più frequente dopo SOV. Le lingue SVO sono VO per definizione (il verbo precede l'oggetto), quindi ci aspetteremmo sulla base degli universali 3 e 4 che abbiano preposizioni. E infatti, molte lingue SVO paradigmatiche hanno proprio preposizioni come tipo dominante di adposizione: lingue romanze come italiano ("di Marco", "a Roma", "da Napoli"), francese ("de Jean", "à Paris"), spagnolo ("de Juan", "a Madrid"), portoghese ("de João", "a Lisboa"); lingue germaniche come inglese ("of John", "to London"), tedesco standard ("von Hans", "nach Berlin"), olandese ("van Jan", "naar Amsterdam"), svedese ("av Johan", "till Stockholm"); e molte altre lingue indoeuropee e non indoeuropee con ordine SVO.

Consideriamo l'italiano come caso paradigmatico di lingua SVO con preposizioni. L'ordine basico dei costituenti in una frase transitiva semplice è soggetto-verbo-oggetto: "Marco legge il libro", dove "Marco" è il soggetto (S), "legge" è il verbo (V), e "il libro" è l'oggetto diretto (O), producendo la sequenza SVO. Le adposizioni sono sistematicamente preposizioni che precedono il loro complemento: "di Marco" (genitivo/possessivo), "a Roma" (direzionale/locativo), "con Maria" (comitativo), "per te" (benefattivo), "su tavolo" (locativo superiore), "in casa" (locativo interno), "da Napoli" (origine), "tra amici" (locativo mediativo). In tutti questi casi, l'adposizione precede il nome, configurando l'ordine Preposizione-Nome, che è armonico con l'ordine generale Verbo-Oggetto della lingua.

Inoltre, seguendo l'Universale 2 discusso precedentemente, dato che l'italiano ha preposizioni, ci aspettiamo che abbia anche genitivo postnominale. E infatti, le costruzioni genitive dell'italiano seguono esattamente questo pattern: "il libro di Marco", "la casa di Maria", "l'auto di Giovanni", dove il modificatore genitivo (introdotto dalla preposizione "di") segue sempre il nome testa. Questo si allinea perfettamente con le previsioni derivanti dall'armonia direzionale: l'italiano è una lingua head-initial nel dominio verbale (V prima di O) e coerentemente head-initial nel dominio adposizionale (P prima di N) e nel dominio della modificazione nominale (N prima di Gen).

Tuttavia, la situazione empirica è più complessa e sfumata di quanto sembri a prima vista da questa caratterizzazione idealizzata, perché esistono lingue classificate come SVO che mostrano pattern misti o che deviano parzialmente da queste correlazioni armoniche, riflettendo il fatto importante che il tipo SVO è in qualche senso intermedio o transizionale tra il tipo fortemente head-initial VSO e il tipo fortemente head-final SOV. Le lingue SVO

mostrano maggiore variabilità e "mescolanza" di proprietà tipologiche rispetto alle lingue SOV e VSO, suggerendo che rappresentano una zona di maggiore instabilità o flessibilità parametrica nello spazio tipologico.

Per esempio, alcune lingue SVO mostrano ordini misti per quanto riguarda i modificatori nominali, con alcuni modificatori che precedono il nome e altri che lo seguono, in violazione apparente dell'armonia direzionale pura. L'inglese, pur essendo chiaramente SVO con preposizioni, ha aggettivi attributivi prenominali ("the red book", "the beautiful house"), che sembrano costituire un'eccezione al pattern head-initial, dato che il nome è la testa del sintagma nominale e gli aggettivi sono modificatori. Tuttavia, va notato che l'inglese ha genitivi prevalentemente postnominali quando espressi con la preposizione "of" ("the book of John"), anche se ha anche un genitivo sassone prenominale con il marcatore 's ("John's book"), mostrando quindi variazione interna.

Il cinese mandarino rappresenta un caso particolarmente interessante e teoricamente problematico. Il cinese è tipicamente classificato come SVO sulla base dell'ordine dei costituenti maggiori in frasi transitive semplici: "Wǒ kàn shū" (Io leggere libro = "Io leggo il/un libro"), con ordine chiaramente SVO. Tuttavia, il cinese mostra una mescolanza notevole di proprietà tipologicamente associate a lingue OV: ha modificatori nominali prevalentemente prenominali inclusi aggettivi ("hóng shū" = rosso libro = "libro rosso"), genitivi ("Zhāngsān de shū" = Zhangsan PARTICELLA libro = "il libro di Zhangsan"), e clausole relative ("wǒ mǎi de shū" = io comprare PARTICELLA libro = "il libro che ho comprato"), tutti precedenti il nome testa. Inoltre, il cinese ha sia preposizioni sia postposizioni, mostrando ancora una volta pattern misto.

Questa configurazione apparentemente contraddittoria del cinese ha generato un dibattito considerevole e prolungato nella letteratura tipologica. Alcune analisi hanno proposto che il cinese mandarino, nonostante l'ordine superficiale SVO nelle frasi principali, mostri caratteristiche strutturali sottostanti di tipo OV in molti domini della grammatica, suggerendo che potrebbe rappresentare una lingua in transizione storica da un tipo OV più antico verso un tipo VO più moderno, o viceversa. Altri studiosi hanno osservato che le costruzioni di modificazione nominale del cinese sono strutturalmente diverse da quelle delle tipiche lingue SVO: piuttosto che essere vere frasi subordinate con struttura clausale completa, molte costruzioni modificatrici del cinese sono più simili a strutture di nominalizzazione o a costruzioni partecipiali, più vicine morfosintatticamente agli aggettivi che alle clausole finite, il che potrebbe spiegare la loro posizione prenominale.

L'Universale 24 introduce una correlazione addizionale e importante tra l'ordine relativo di soggetto, oggetto e verbo e la posizione degli aggettivi qualificativi o descrittivi rispetto al nome testa nel sintagma nominale. Nella sua formulazione originale, esso afferma che se la posizione relativa di soggetto e oggetto è SOV o OSV (cioè se O precede V nella serializzazione superficiale), allora la lingua ha sempre o quasi sempre aggettivi prenominali, ovvero aggettivi che precedono sistematicamente il nome che modificano. In altre parole, le lingue che collocano l'oggetto prima del verbo a livello della frase transitiva tendono sistematicamente e robustatamente a collocare anche gli aggettivi prima del nome a livello del sintagma nominale.

Formalmente, possiamo esprimere questo universale come: P(Aggettivi prenominali | OV) ≈ 0,85-0,90, dove OV denota l'insieme di lingue con ordine OV (principalmente SOV ma includendo anche OSV e OVS). Questo universale cattura ancora una volta la tendenza generale all'armonia direzionale: se i complementi verbali precedono il verbo manifestando l'ordine O-V e quindi una configurazione head-final a livello della proiezione verbale, anche i modificatori nominali tenderanno a precedere il nome manifestando l'ordine Aggettivo-Nome e quindi una configurazione head-final consistente a livello della proiezione nominale.

Lingue prototipiche SOV come il turco, il giapponese, il coreano, l'hindi, il tamil, e il basco mostrano infatti aggettivi prenominali, confermando empiricamente la previsione dell'universale con notevole consistenza. In turco: "kırmızı kitap" (rosso libro = "libro rosso"), "büyük ev" (grande casa = "casa grande"), dove l'aggettivo precede sempre il nome. In giapponese: "akai hon" (rosso libro = "libro rosso"), "ōkii ie" (grande casa = "casa grande"), con la stessa configurazione prenominale. In hindi: "lāl kitāb" (rosso libro = "libro rosso"), "baṛā ghar" (grande casa = "casa grande"), ancora con aggettivi prenominali. La consistenza di questo pattern attraverso lingue geneticamente non correlate e geograficamente distanti fornisce evidenza forte per la validità dell'universale come generalizzazione tipologica robusta.

Al contrario, le lingue con ordine VO mostrano maggiore variabilità e minor uniformità nella posizione degli aggettivi, riflettendo ancora una volta che il tipo VO (e specialmente il sottotipo SVO) è meno armonicamente coerente del tipo OV. Alcune lingue VO, in particolare molte lingue VSO come il gallese, l'irlandese, il tagalog e varie lingue celtiche e austronesiane, hanno aggettivi postnominali in perfetta armonia con l'ordine generale head-initial o testa-complemento. In gallese: "llyfr coch" (libro rosso = "libro rosso"), "tŷ mawr" (casa grande = "casa grande"), dove l'aggettivo segue il nome. In irlandese: "leabhar dearg" (libro rosso = "libro rosso"), con ordine similare Nome-Aggettivo.

Altre lingue VO, tuttavia, in particolare molte lingue SVO come l'inglese, il tedesco, l'olandese, e varie lingue germaniche, hanno aggettivi prevalentemente o esclusivamente prenominali, costituendo apparenti controesempi all'armonia direzionale stretta. In inglese: "red book" (libro rosso), "big house" (casa grande), con ordine Aggettivo-Nome. In tedesco: "rotes Buch" (libro rosso), "großes Haus" (casa grande), ancora con aggettivi prenominali. Questa deviazione dall'armonia suggerisce che la posizione degli aggettivi è influenzata da molteplici fattori che possono competere o interagire in modi complessi.

La posizione degli aggettivi è infatti influenzata da una molteplicità di fattori linguistici oltre all'armonia direzionale basica, inclusi: primo, il tipo semantico di aggettivo, con distinzioni tra aggettivi descrittivi versus classificatori, aggettivi valutativo-soggettivi versus oggettivo-descrittivi, aggettivi restrittivi versus non restrittivi; secondo, la complessità sintattica del modificatore, dove aggettivi semplici e aggettivi con complementi o modificatori propri possono avere posizioni diverse; terzo, fattori di peso sintattico e preferenze di ramificazione; quarto, influenze storiche e diacroniche, dove la posizione degli aggettivi può riflettere stadi precedenti dell'evoluzione della lingua.

L'analisi dettagliata dell'inglese rivela pattern più complessi che riflettono tensioni tra diverse pressioni strutturali. Gli aggettivi attributivi semplici sono rigidamente prenominali ("the beautiful house", *"the house beautiful"), ma aggettivi con complementi o modificatori

pesanti tendono a essere postnominali ("a house [beautiful beyond description]", "something [interesting to read]"). Inoltre, certi aggettivi di origine latina hanno posizione postnominale in espressioni fisse o tecniche ("attorney general", "court martial", "body politic"), riflettendo l'influsso storico del francese normanno che aveva aggettivi postnominali.

Le lingue romanze mostrano variazione interessante. Il francese, lo spagnolo e l'italiano hanno generalmente aggettivi postnominali come ordine basico, in armonia con l'ordine head-initial: francese "livre rouge" (libro rosso), "maison grande" (casa grande); spagnolo "libro rojo", "casa grande"; italiano "libro rosso", "casa grande". Tuttavia, alcune sottoclassi di aggettivi, tipicamente quelli molto frequenti, brevi, e soggettivi/valutativi, possono o devono precedere il nome: francese "bon livre" (buon libro), "grand homme" (grande uomo); spagnolo "buen libro", "gran hombre"; italiano "buon libro", "grand'uomo" (arcaico). Questa variazione suggerisce che la posizione dell'aggettivo è soggetta a fattori lessicali e semantici oltre che a principi sintattici generali.

Le Clausole Relative e la Loro Posizione: Un'Analisi Approfondita dell'Universale 24 e delle Sue Ramificazioni Teoriche

Le clausole relative rappresentano indiscutibilmente uno dei terreni di analisi più ricchi, complessi e teoricamente illuminanti per lo studio degli universali linguistici e delle correlazioni tipologiche sistematiche. Una clausola relativa è una clausola subordinata o dipendente che modifica un sintagma nominale (la cosiddetta "testa" della relativa) fornendo informazioni aggiuntive, restrittive o non restrittive, su di esso. Le clausole relative sono strutture sintatticamente complesse che coinvolgono tipicamente un elemento relativi (pronome relativo o complementatore relativo) e una lacuna o gap strutturale corrispondente a un argomento o aggiunto mancante che viene identificato con la testa nominale.

Per esempio, in italiano la frase "il libro che ho letto ieri" contiene una clausola relativa ("che ho letto ieri") che modifica il nome testa "libro". La struttura della clausola relativa può essere analizzata come contenente un complementatore "che" e una struttura clausale con una posizione oggetto vuota o gap: "[che [io ho letto ___ ieri]]", dove il gap (___) corrisponde all'oggetto del verbo "letto" ed è co-indicizzato con la testa "libro". La posizione delle clausole relative rispetto alla testa nominale mostra una variazione tipologica significativa e strutturata, e, cosa teoricamente cruciale per la nostra discussione, questa posizione correla sistematicamente con altri aspetti dell'ordine delle parole secondo pattern implicazionali identificati originariamente da Greenberg e successivamente raffinati.

L'Universale 24, nella sua formulazione originale di Greenberg che ha subito varie interpretazioni e raffinamenti nella letteratura successiva, afferma nella sua forma più diretta che se la posizione relativa di soggetto e oggetto è SOV (ovvero se l'oggetto precede il verbo), allora la lingua ha sempre o quasi sempre clausole relative prenominali, cioè clausole relative che precedono sistematicamente il nome testa che modificano. In altre parole, le lingue in cui l'oggetto precede il verbo nella frase principale transitiva tendono sistematicamente e robustatamente a collocare le clausole relative prima del nome che

modificano, manifestando l'ordine [Clausola Relativa] Nome piuttosto che Nome [Clausola Relativa].

Formalmente, possiamo esprimere questo universale come una probabilità condizionale: P(RC prenominali | SOV) ≈ 0,85-0,90, dove RC denota "relative clauses" (clausole relative). Questo valore probabilistico indica che, data una lingua con ordine SOV, la probabilità che abbia clausole relative prenominali è molto alta, dell'ordine del 85-90% secondo le stime più affidabili basate su campioni tipologici ampi e diversificati. La deviazione sostanziale da 0,50 (che rappresenterebbe assenza di correlazione) dimostra una forte associazione statistica.

Questo pattern si osserva con notevole regolarità e consistenza in numerose lingue prototipicamente SOV appartenenti a diverse famiglie genetiche e distribuite in diverse aree geografiche: lingue altaiche come turco, azero, uzbeko, kazako, kirghiso, mongolo, buriato, tunguso; lingue dravidiche come tamil, telugu, kannada, malayalam; lingue sino-tibetane come birmano, tibetano; lingue uralico-finniche come finlandese, estone, ungherese; lingue isolate come giapponese, coreano, basco; lingue indo-iraniane orientali come hindi, urdu, punjabi, bengali; e molte altre lingue sparse geograficamente ma convergenti tipologicamente.

Consideriamo un'analisi dettagliata con esempi specifici dal turco, una lingua SOV prototipica con clausole relative prenominali particolarmente chiare e regolari. La frase italiana "il libro che ho letto" corrisponde in turco a "okuduğum kitap", che può essere glosata morfema per morfema come "leggere-PASSATO-POSSESSIVO.1SG libro". Analizziamo la struttura: "oku-" è la radice verbale "leggere"; "-duğ-" è un suffisso di nominalizzazione/participio passato; "-um" è un suffisso possessivo di prima persona singolare che marca l'agente della nominalizzazione; "kitap" significa "libro". La clausola relativa nominalizzata "okuduğum" precede chiaramente e categoricamente la testa nominale "kitap".

La struttura interna della clausola relativa turca riflette l'ordine SOV caratteristico della lingua. In una frase principale transitiva come "ho letto il libro", il turco usa l'ordine "(ben) kitabı okudum", letteralmente "io libro-ACCUSATIVO leggere-PASSATO-1SG", con l'oggetto "kitabı" (libro marcato con il caso accusativo) che precede il verbo "okudum" (lessi). Quando questa struttura transitiva viene nominalizzata per formare una clausola relativa, mantiene la stessa configurazione strutturale OV, semplicemente incorporando il tutto come modificatore prenominale della testa attraverso il processo di nominalizzazione participiale.

Consideriamo un esempio più complesso per illustrare la robustezza del pattern. La frase "il libro che l'uomo ha dato alla donna" in turco è "adamın kadına verdiği kitap", analizzabile come: "adam-ın" (uomo-GENITIVO), "kadın-a" (donna-DATIVO), "ver-diğ-i" (dare-NOMINALIZZATORE-POSSESSIVO.3SG), "kitap" (libro). Tutta la clausola relativa complessa "adamın kadına verdiği" precede il nome "kitap", mantenendo internamente l'ordine SOV: il soggetto "adamın" precede l'oggetto indiretto "kadına" che precede il verbo nominalizzato "verdiği". Questa consistenza strutturale multi-livello dimostra l'armonia tipologica profonda della lingua.

Il giapponese, un'altra lingua SOV prototipica, mostra pattern strettamente paralleli. La frase "il libro che ho letto" si rende in giapponese come "watashi ga yonda hon", analizzabile come "io-NOMINATIVO leggere-PASSATO libro". La clausola relativa "watashi ga yonda" (io

leggere-PASSATO) precede categoricamente la testa "hon" (libro). Notare che il giapponese, a differenza del turco che usa nominalizzazione esplicita con morfologia derivazionale, usa semplicemente una forma verbale finita (il passato "yonda") senza marcatura speciale per introdurre la clausola relativa, riflettendo una strategia di relativizzazione diversa ma mantenendo la stessa posizione prenominale.

Il pattern opposto e speculare si osserva sistematicamente nelle lingue VO, in particolare quelle SVO come l'italiano, l'inglese, il francese, lo spagnolo e la maggior parte delle lingue romanze e germaniche. In queste lingue le clausole relative seguono tipicamente e preferenzialmente il nome che modificano, manifestando l'ordine Nome [Clausola Relativa] in armonia con l'ordine generale head-initial o testa-complemento. Formalmente: P(RC postnominali | VO) ≈ 0,75-0,85, indicando una correlazione forte anche se leggermente meno robusta di quella nelle lingue OV, riflettendo la maggiore variabilità tipologica delle lingue VO.

In italiano, come già osservato, le clausole relative sono sistematicamente postnominali: "il libro [che ho letto]", "la persona [che è arrivata]", "lo studente [a cui ho parlato]", dove le clausole relative in parentesi quadre seguono sempre il nome testa. La struttura interna della clausola relativa riflette l'ordine SVO caratteristico dell'italiano: dopo il pronome relativo o complementatore "che"/"cui" che introduce la clausola, segue una struttura con ordine soggetto-verbo(-oggetto) quando il gap non è in posizione soggetto.

L'inglese mostra pattern paralleli con clausole relative postnominali: "the book [that I read]", "the person [who arrived]", "the student [to whom I spoke]". La struttura interna mantiene l'ordine SVO basico: "that I read ___" ha ordine soggetto-verbo con gap in posizione oggetto, mentre "who ___ arrived" ha gap in posizione soggetto seguito dal verbo, mantenendo la posizione V dopo S.

Il francese: "le livre [que j'ai lu]", "la personne [qui est arrivée]", "l'étudiant [à qui j'ai parlé]". Lo spagnolo: "el libro [que leí]", "la persona [que llegó]", "el estudiante [a quien hablé]". In tutti questi casi, la clausola relativa segue il nome in armonia con l'ordine generale VO della lingua, e la struttura interna mantiene l'ordine VO/SVO caratteristico.

Questa correlazione robusta tra ordine OV/VO e posizione delle clausole relative non è assoluta ma statistica, come già enfatizzato, e le eccezioni apparenti o i casi problematici offrono opportunità preziose e teoricamente illuminanti per raffinare la nostra comprensione dei fattori multipli in gioco. Una complicazione importante e ampiamente discussa nella letteratura deriva dal fatto che alcune lingue, come il cinese mandarino menzionato precedentemente, sono classificate come SVO per quanto riguarda l'ordine dei costituenti maggiori della frase transitiva basica, ma mostrano clausole relative prenominali, apparentemente violando o costituendo un controesempio all'Universale 24.

La frase "il libro che ho letto" in cinese mandarino si rende come "wǒ dú de shū", analizzabile morfema per morfema come "io leggere PARTICELLA libro". Analizziamo la struttura: "wǒ" è il pronome di prima persona singolare "io"; "dú" è il verbo "leggere"; "de" è una particella di modificazione nominalizzante o relativizzante; "shū" significa "libro". La clausola relativa "wǒ dú" (io leggere) precede chiaramente la testa "shū" (libro), manifestando ordine prenominale simile alle lingue SOV.

Questo apparente controesempio ha generato un dibattito teorico considerevole e prolungato nella letteratura tipologica e sintattica. Diverse analisi e proposte sono state avanzate per spiegare o rianalizzare questo caso problematico. Primo, alcune analisi hanno proposto che il cinese mandarino, nonostante l'ordine SVO superficiale nelle frasi principali monoclausal, mostri caratteristiche sottostanti di tipo OV in molti domini della grammatica oltre alle clausole relative, inclusa la presenza di postposizioni coesistenti con preposizioni (il cinese ha entrambi i tipi di adposizioni), la tendenza generale a collocare modificatori prima delle teste nel dominio nominale (aggettivi prenominali, genitivi prenominali, numerali prenominali), e altre proprietà sintattiche che suggeriscono un sistema misto o in transizione tipologica.

Secondo questa prospettiva analitica, il cinese potrebbe essere concettualizzato non come una pura lingua SVO del tipo romanzo-germanico, ma come una lingua con ordine SVO innovativo derivante da un più antico ordine SOV, che mantiene però relitti o persistenze del tipo OV più antico in vari domini grammaticali incluse le clausole relative. Evidenza diacronica dal cinese arcaico e dal cinese classico suggerisce che stadi più antichi della lingua potrebbero aver avuto ordine diverso, possibilmente con maggiore prominenza di pattern OV.

Terzo, altri studiosi hanno osservato che le clausole relative del cinese sono strutturalmente diverse e morfosintatticamente non parallele a quelle delle tipiche lingue SVO indoeuropee: piuttosto che essere vere clausole subordinate con struttura clausale completa e un pronome relativo o complementatore che introduce una frase finita, le relative del cinese sono morfologicamente più simili a strutture di nominalizzazione o a costruzioni partecipiali prenominali, avvicinandosi morfosintatticamente più agli aggettivi modificatori che alle clausole subordinate complete. La particella "de" che marca queste costruzioni è multifunzionale e appare in vari contesti di modificazione nominale, non solo nelle relative, supportando l'idea che queste costruzioni siano nominali piuttosto che verbali nella loro natura categoriale profonda.

Quarto, alcune analisi sintattiche formali hanno proposto che le cosiddette "clausole relative" prenominali del cinese non siano derivate per movimento di un operatore relativo come nelle lingue indoeuropee, ma siano semplicemente clausole aggiunte o appositive in posizione di modificatore, senza la struttura operatore-variabile caratteristica delle vere relative. Questa differenza strutturale profonda potrebbe spiegare perché seguono principi di collocazione diversi.

Un altro aspetto degno di nota riguarda le lingue con ordine VSO, che rappresentano circa il 9% delle lingue del mondo secondo le stime tipologiche contemporanee. L'Universale 24 nella sua formulazione originale non fa previsioni specifiche esplicite per queste lingue, dato che si concentra sulla posizione relativa di S e O piuttosto che sulla posizione assoluta di V. Tuttavia, dato che le lingue VSO sono per definizione lingue VO (il verbo precede l'oggetto), l'estensione naturale dell'universale prevederebbe clausole relative postnominali per consistenza con l'ordine head-initial.

La verifica empirica mostra che molte lingue celtiche con ordine VSO, come l'irlandese, il gallese, il bretone e lo scozzese gaelico, hanno effettivamente clausole relative postnominali, in armonia con il loro carattere generale head-initial. In gallese, una lingua

VSO ben studiata, la struttura basica è verbo-soggetto-oggetto: "Gwelodd Siôn ddraig" (vedere-PASSATO Siôn drago = "Siôn vide un drago"), con ordine VSO. Le clausole relative seguono il nome: "y ddraig [a welodd Siôn]" (il drago [PARTICELLA vedere-PASSATO Siôn] = "il drago che Siôn vide"), con la clausola relativa in posizione postnominale.

In irlandese: "an leabhar [a léigh mé]" (il libro [PARTICELLA leggere-PASSATO io] = "il libro che lessi"), ancora con clausola relativa postnominale. Tuttavia, va notato che alcune lingue VSO mostrano pattern più complessi, con diverse strategie di relativizzazione disponibili o con la posizione della clausola relativa che dipende da fattori come il tipo di nome relativizzato (nomi propri versus nomi comuni, nomi definiti versus indefiniti) o la funzione sintattica del gap all'interno della relativa (relativizzazione di soggetto versus oggetto versus obliquo).

La spiegazione teorica più influente e ampiamente adottata per la correlazione tra ordine OV/VO e posizione delle clausole relative fa ancora una volta appello al principio generale di armonia direzionale o di branching consistente, ora applicato specificamente al dominio della modificazione clausale complessa. L'idea centrale è che le lingue mostrino una preferenza sistematica per collocare il materiale sintattico pesante o strutturalmente complesso (come una clausola relativa, che è tipicamente più lunga in termini di numero di parole e strutturalmente più elaborata in termini di profondità sintattica rispetto a un semplice aggettivo o sintagma adposizionale) nella stessa direzione relativa alla testa in cui collocano altro materiale dipendente o complementare.

Le lingue OV, che sistematicamente collocano i complementi verbali prima delle loro teste verbali manifestando l'ordine Oggetto-Verbo, tenderanno per armonia strutturale a collocare anche le clausole relative (che sono un tipo complesso di modificatore del nome) prima della testa nominale che modificano, manifestando l'ordine [Clausola Relativa]-Nome. Le lingue VO, che collocano i complementi verbali dopo le teste verbali manifestando l'ordine Verbo-Oggetto, tenderanno simmetricamente a collocare le clausole relative dopo la testa nominale, manifestando l'ordine Nome-[Clausola Relativa]. In entrambi i casi, la direzione di ramificazione (branching) è consistente: o a sinistra (left-branching) per le lingue OV, o a destra (right-branching) per le lingue VO.

Un raffinamento teorico importante di questa analisi, sviluppato ampiamente nella letteratura psicolinguistica e sintattica, considera il concetto di ramificazione (branching) e peso sintattico in termini più quantitativi e graduali piuttosto che puramente qualitativi. È stato osservato attraverso numerosi studi psicolinguistici che molte lingue mostrano una preferenza generale per la ramificazione a peso crescente, ovvero una tendenza a collocare costituenti sintatticamente più pesanti (più lunghi, più complessi, più profondamente annidati) verso la fine della frase o del sintagma piuttosto che all'inizio o nel mezzo.

Questa preferenza può essere motivata da principi di processamento e produzione linguistica: costituenti pesanti collocati alla fine permettono al parlante di pianificare incrementalmente la struttura mentre procede, e permettono all'ascoltatore di processare incrementalmente l'informazione senza dover tenere in memoria working strutture complesse non risolte. Le clausole relative, essendo costituenti intrinsecamente complessi e potenzialmente molto lunghi (possono contenere frasi complete con argomenti multipli,

modificatori, e subordinate ulteriori), si prestano particolarmente bene a testare questa generalizzazione sul peso sintattico.

Le lingue fortemente right-branching come l'inglese collocano sistematicamente le clausole relative dopo il nome, permettendo la ramificazione a destra e facilitando il processamento incrementale. Consideriamo la differenza tra "the book [that I read yesterday in the library during my lunch break]" (perfettamente naturale, con clausola relativa potenzialmente lunga a destra) versus "*[that I read yesterday in the library during my lunch break] book" (totalmente agrammaticale in inglese, con clausola relativa pesante a sinistra).

Le lingue fortemente left-branching come il giapponese collocano sistematicamente le clausole relative prima del nome, manifestando ramificazione a sinistra consistente ma accettando facilmente clausole relative molto lunghe e complesse in questa posizione. Il giapponese permette costruzioni come "[watashi ga kinō toshokan de hirugohan no aida ni yonda] hon" (letteralmente "[io NOMINATIVO ieri biblioteca in pranzo GENITIVO durante leggere-PASSATO] libro"), con una clausola relativa lunga e complessa in posizione prenominale, senza problemi di accettabilità o processamento.

Tuttavia, la realtà tipologica è più sfumata e complessa di quanto suggerisca un'applicazione rigida e uniforme del principio di ramificazione consistente. Molte lingue mostrano ramificazione mista, con alcuni tipi di costituenti che ramificano a destra e altri a sinistra, riflettendo che diverse pressioni strutturali e di processamento possono operare differentemente in diversi domini grammaticali.

L'inglese, per esempio, è prevalentemente right-branching per quanto riguarda le clausole relative postnominali, i complementi verbali postverbali (V-O order), le clausole complemento postverbali ("I think [that John left]"), le frasi preposizionali che seguono il nome ("the book [on the table]"), ma mostra left-branching nel dominio della modificazione nominale mediante aggettivi semplici, che precedono rigidamente il nome ("the beautiful house", *"the house beautiful"). Questa asimmetria suggerisce che aggettivi semplici e clausole relative, pur essendo entrambi modificatori nominali, sono soggetti a principi diversi di collocazione.

Il cinese mandarino, come già discusso, mostra il pattern opposto e complementare: è prevalentemente left-branching per quanto riguarda i modificatori nominali (aggettivi prenominali, genitivi prenominali, clausole relative prenominali, numerali prenominali), ma right-branching per quanto riguarda i complementi verbali (ordine V-O), mostrando una dissociazione netta tra il dominio nominale e il dominio verbale che viola l'armonia direzionale semplice.

Questa variabilità inter-linguistica e intra-linguistica ha portato alcuni ricercatori a proporre gerarchie di applicazione delle tendenze armoniche, dove certi domini grammaticali sono più resistenti al cambiamento e più fortemente influenzati dall'armonia direzionale, mentre altri domini sono più flessibili e più suscettibili a pressioni indipendenti. Secondo queste analisi gerarchiche, le clausole relative si collocano in una posizione intermedia nella gerarchia di armonia: sono fortemente influenzate dall'ordine OV/VO ma non rigidamente e deterministicamente determinate da esso, permettendo eccezioni motivate da fattori come il peso sintattico specifico, la complessità strutturale, le pressioni comunicative legate all'elaborazione e alla comprensione, o relitti storici di stadi precedenti della lingua.

Alcune proposte teoriche hanno suggerito scale implicazionali per la coerenza direzionale nei diversi domini. Una possibile gerarchia, basata sull'osservazione tipologica, potrebbe essere: Adposizioni > Genitivi > Clausole Relative > Aggettivi > Altri modificatori, dove > indica "più armonico di" o "più fortemente correlato con l'ordine V-O di". Secondo questa gerarchia, se una lingua viola l'armonia direzionale in un dominio più alto (come avendo postposizioni in una lingua VO), violerà necessariamente anche l'armonia nei domini più bassi. Ma se viola l'armonia in un dominio più basso (come avendo aggettivi prenominali in una lingua VO), può ancora mantenere l'armonia nei domini più alti.

Gli Universali e i Casi Grammaticali: Un'Analisi Dettagliata di Genitivi, Adposizioni e Marcatura dei Ruoli Semantici

Un'area particolarmente ricca di universali implicazionali e correlazioni tipologiche riguarda l'espressione morfologica e sintattica dei casi grammaticali e dei ruoli semantici associati agli argomenti nominali. I casi grammaticali sono marcature morfologiche sui nomi (e sui loro modificatori e determinanti in lingue con concordanza) che indicano la funzione sintattica che il nome svolge nella frase (soggetto, oggetto diretto, oggetto indiretto, etc.) o il ruolo semantico associato (agente, paziente, beneficiario, strumento, locazione, etc.). Alcune lingue, come il latino classico, il russo moderno, il finlandese, il turco, il giapponese e molte altre lingue con morfologia nominale ricca, hanno sistemi di casi molto elaborati con molte distinzioni morfologiche esplicite che codificano fino a quindici o venti casi diversi, mentre altre lingue, come l'inglese moderno, il cinese mandarino, il francese moderno e molte lingue isolanti o analitiche, hanno perso quasi completamente o completamente la flessione nominale casuale e ricorrono invece principalmente all'ordine fisso delle parole e alle adposizioni (preposizioni o postposizioni) per indicare le relazioni grammaticali e semantiche.

L'Universale 41 stabilisce una correlazione importante e robusta tra il tipo di sistema di marcatura del caso e l'ordine dei costituenti maggiori. Nella sua formulazione originale, esso afferma che se in una lingua la distinzione nominativo-accusativo (la distinzione fondamentale tra il caso del soggetto e il caso dell'oggetto diretto in lingue con allineamento nominativo-accusativo) è marcata morfologicamente solo nel sistema pronominale e non nel sistema dei nomi comuni, allora la lingua è quasi certamente di tipo VO (verbo precede oggetto).

Questo universale identifica un'asimmetria tipologica interessante e teoricamente significativa: le lingue che hanno una marcatura ridotta o atrofizzata del caso, limitata morfologicamente ai pronomi senza estendersi ai nomi comuni, tendono sistematicamente ad essere lingue VO, suggerendo che l'ordine VO (verbo prima dell'oggetto con conseguente ordine relativamente fisso SVO o VSO) permette, facilita o forse addirittura favorisce una riduzione della marcatura morfologica esplicita del caso, perché la posizione strutturale fissa del costituente nominale può indicare sufficientemente e non ambiguamente la sua funzione sintattica senza bisogno di marcatura morfologica ridondante.

Al contrario, le lingue OV, specialmente quelle con ordine dei costituenti relativamente flessibile o libero che permette scrambling o topicalizzazione, tendono a mantenere sistemi di marcatura del caso più robusti ed elaborati, perché la marcatura morfologica esplicita diventa funzionalmente necessaria per disambiguare le funzioni sintattiche quando l'ordine delle parole non è rigido e affidabile.

Formalmente, possiamo esprimere questo universale come: P(VO | Caso solo pronominale) ≈ 0,90-0,95, ovvero la probabilità che una lingua sia VO dato che ha marcatura del caso limitata al sistema pronominale è molto alta. Equivalentemente: P(Caso solo pronominale | OV) ≈ 0,05-0,10, ovvero la probabilità che una lingua abbia marcatura del caso solo pronominale dato che è OV è molto bassa.

L'italiano moderno fornisce un esempio particolarmente chiaro e paradigmatico di questo pattern tipologico. L'italiano è una lingua SVO (quindi VO) con una marcatura del caso morfologica estremamente limitata e atrofizzata. La distinzione nominativo-accusativo, che è la distinzione casuale più fondamentale e universalmente attestata nelle lingue con allineamento nominativo-accusativo, è visibile morfologicamente solo in alcuni pronomi personali tonici e clitici: "egli/lui" (nominativo forma tonica soggetto / forma obliqua usata come accusativo), "ella/lei" (nominativo femminile / accusativo femminile), "io/me" (nominativo prima persona / accusativo/obliquo prima persona), "tu/te" (nominativo seconda persona / accusativo seconda persona), "noi" (invariabile), "voi" (invariabile).

Per i nomi comuni, al contrario, non esiste alcuna marcatura morfologica distinguibile che differentii il soggetto dall'oggetto: "Marco vede Lucia" e "Lucia vede Marco" hanno significati radicalmente diversi derivanti esclusivamente dall'ordine delle parole, con il primo nome nella sequenza interpretato grammaticalmente come soggetto e il secondo come oggetto secondo la grammatica configurazionale rigida SVO dell'italiano. Questa dipendenza totale dall'ordine delle parole per indicare le funzioni sintattiche è resa possibile e non problematica dal fatto che l'italiano ha un ordine basico SVO relativamente rigido e non permette scrambling libero degli argomenti maggiori.

Questo pattern si allinea perfettamente con la previsione dell'Universale 41: l'italiano ha marcatura del caso limitata al sistema pronominale ed è effettivamente una lingua VO (specificamente SVO). Il francese moderno mostra un pattern ancora più estremo di atrofia casuale: anche i pronomi soggetto sono quasi completamente invariabili o indistinguibili foneticamente ("je/me", "tu/te", ma "il/le" suonano identici in molte varietà parlate), rendendo l'ordine delle parole ancora più cruciale per l'interpretazione sintattica.

Al contrario, lingue SOV prototipiche come il turco, il giapponese, il coreano, l'hindi e il tamil hanno tipicamente sistemi di marcatura del caso morfologica molto più sviluppati, robusti ed elaborati, estesi non solo ai pronomi ma anche ai nomi comuni. Il turco, per esempio, marca morfologicamente l'oggetto diretto definito con il suffisso di accusativo -ı/i/u/ü (con variazione per armonia vocalica secondo le regole fonologiche della lingua): "Kitabı gördüm" significa "ho visto il libro" o "vidi il libro", dove "kitab" (libro, radice nominale) riceve il suffisso accusativo -ı per indicare esplicitamente che è l'oggetto diretto del verbo "gördüm" (vidi).

Il soggetto, al contrario, appare tipicamente senza marcatura esplicita quando è al caso nominativo, che è il caso non marcato o zero: "Kitap burada" (libro qui = "il libro è qui"), dove "kitap" appare senza suffisso casuale. Tuttavia, il turco marca morfologicamente anche molti

altri casi oltre al nominativo e accusativo, inclusi genitivo (-ın/in/un/ün), dativo (-a/e), locativo (-da/de), ablativo (-dan/den), strumentale-comitativo (-la/le), creando un sistema ricco di sei casi morfologicamente distinti più varie postposizioni che governano casi specifici.

Questo sistema elaborato di marcatura del caso permette al turco una certa flessibilità nell'ordine delle parole per scopi pragmatici o enfatici: l'ordine basico è SOV, ma permutazioni come OSV o topicalizzazioni sono possibili perché la funzione sintattica è codificata morfologicamente oltre che posizionalmente. Per esempio, "Kitabı gördüm ben" (libro-ACC vidi io = "il libro, l'ho visto io") permuta l'oggetto in posizione iniziale per topicalizzazione, ma la marcatura accusativa -ı garantisce che "kitabı" sia interpretato come oggetto nonostante la posizione non canonica.

Il giapponese manifesta un pattern simile ma con un sistema di particelle postposizionali piuttosto che suffissi flessivi. Le funzioni sintattiche sono indicate da particelle invariabili che seguono i sintagmi nominali: "ga" marca il soggetto/nominativo, "o" marca l'oggetto diretto/accusativo, "ni" marca il dativo/locativo, "de" marca lo strumentale/locativo, "no" marca il genitivo, "to" marca il comitativo, etc. Per esempio: "Tarō ga hon o yonda" (Taro NOMINATIVO libro ACCUSATIVO leggere-PASSATO = "Taro lesse il/un libro"), dove le particelle "ga" e "o" marcano esplicitamente le funzioni sintattiche.

Questa marcatura esplicita permette al giapponese scrambling relativamente libero: "Hon o Tarō ga yonda" (libro ACCUSATIVO Taro NOMINATIVO leggere-PASSATO) permuta l'oggetto in posizione iniziale ma mantiene interpretazione non ambigua grazie alle particelle. L'ordine basico preferito rimane SOV, ma deviazioni sono grammaticali e funzionalmente motivate da considerazioni pragmatiche.

Il rapporto sistematico tra genitivi e adposizioni, già ampiamente discusso nell'ambito dell'Universale 2, rivela ulteriori pattern correlazionali e implicazionali che arricchiscono il quadro teorico. Greenberg osserva con precisione empirica che nelle lingue con preposizioni (adposizioni che precedono il loro complemento nominale), il genitivo tende robustatamente a seguire il nome governante o testa, generando costruzioni morfosintattiche come "il libro di Marco" in italiano, "the book of John" in inglese, "le livre de Jean" in francese, dove la preposizione "di"/"of"/"de" introduce esplicitamente il genitivo e lo marca come dipendente sintattico e semantico del nome precedente.

Queste costruzioni genitive preposizionali risultano particolarmente trasparenti e facilmente analizzabili dal punto di vista sintattico, perché l'elemento adposizionale (la preposizione genitiva "di" in italiano, "of" in inglese, "de" in francese, spagnolo e portoghese, "von" in tedesco per il genitivo partitivo, oppure la particella genitiva "の" (no) in giapponese, la desinenza "-ın/-in/-un/-ün" in turco) marca esplicitamente e morfologicamente la relazione di dipendenza tra il nome possessore o modificatore (Marco, John, Jean, Mehmet) e il nome posseduto o testa (libro, book, livre, kitap).

Questa marcatura esplicita attraverso preposizioni specializzate contrasta nettamente con le costruzioni genitive delle lingue con postposizioni, dove la relazione genitiva è tipicamente marcata attraverso morfologia flessiva sul nome dipendente stesso (suffissi o desinenze casuali che si agglutinano alla radice nominale), e il genitivo morfologicamente marcato precede la testa nominale senza la mediazione necessaria di elementi adposizionali indipendenti e fonologicamente separati. Nel turco, come già osservato precedentemente in

relazione all'Universale 2, il genitivo è espresso attraverso il suffisso casuale -ın/in/un/ün che si aggiunge alla radice nominale secondo le regole sistematiche e predicibili di armonia vocalica anteriore/posteriore e arrotondata/non arrotondata che caratterizzano la fonologia turca: "Mehmet'in kitabı" (Mehmet-GENITIVO libro-POSSESSIVO.3SG = "il libro di Mehmet"), dove il genitivo morfologicamente marcato "Mehmet'in" precede rigidamente e categoricamente il nome testa "kitap", in perfetta armonia strutturale e direzionale con il pattern generale OV della lingua (l'oggetto precede il verbo nella frase transitiva basica) e con la presenza sistematica di postposizioni piuttosto che preposizioni come categoria grammaticale dominante.

La correlazione statistica robusta e empiricamente ben documentata tra tipo di adposizione (preposizioni versus postposizioni) e posizione del genitivo (postnominale versus prenominale) può essere quantificata con maggiore precisione metodologica utilizzando dati estratti da database tipologici contemporanei di ampia copertura. Se definiamo quattro categorie logicamente possibili nella forma di una tabella di contingenza bidimensionale 2×2: (1) Preposizioni combinate con Genitivo postnominale, rappresentando la configurazione armonica per lingue VO con ordine generale head-initial; (2) Preposizioni combinate con Genitivo prenominale, rappresentando una configurazione disarmonica e teoricamente problematica; (3) Postposizioni combinate con Genitivo postnominale, anch'essa configurazione disarmonica; (4) Postposizioni combinate con Genitivo prenominale, rappresentando la configurazione armonica per lingue OV con ordine generale head-final, l'Universale 2 di Greenberg predice esplicitamente e formalmente che le categorie 1 e 4 (le configurazioni armoniche direzionalmente consistenti) siano densamente popolate nella distribuzione tipologica mondiale, mentre le categorie 2 e 3 (le configurazioni disarmoniche e direzionalmente inconsistenti) siano scarsamente popolate, costituendo al massimo eccezioni marginali che richiedono spiegazioni speciali derivanti da fattori storici, areali o strutturali particolari.

Dati empirici sistematici estratti dal World Atlas of Language Structures (WALS) su un campione metodologicamente controllato e geneticamente diversificato di circa 900 lingue per le quali sono disponibili informazioni complete e affidabili su entrambe le variabili critiche (tipo di adposizione e posizione del genitivo) suggeriscono la seguente distribuzione approssimativa che conferma in modo impressionante e statisticamente robusto le previsioni derivanti dall'Universale 2:

- Categoria 1 (Preposizioni + Genitivo postnominale): circa 420 lingue, rappresentanti il 46,7% del campione totale
- Categoria 4 (Postposizioni + Genitivo prenominale): circa 380 lingue, rappresentanti il 42,2% del campione totale
- Categoria 2 (Preposizioni + Genitivo prenominale): circa 60 lingue, rappresentanti solo il 6,7% del campione totale
- Categoria 3 (Postposizioni + Genitivo postnominale): circa 40 lingue, rappresentanti solo il 4,4% del campione totale

Possiamo calcolare le probabilità condizionali che quantificano la forza dell'implicazione direzionale. La probabilità che una lingua con preposizioni abbia genitivo postnominale è data da: P(Genitivo_postnominale | Preposizioni) = 420 / (420 + 60) = 420 / 480 ≈ 0,875, ovvero circa l'87,5%. Complementarmente, la probabilità che una lingua con postposizioni

abbia genitivo prenominale è: P(Genitivo_prenominale | Postposizioni) = 380 / (380 + 40) = 380 / 420 ≈ 0,905, ovvero circa il 90,5%. Questi valori quantitativi, significativamente e sostanzialmente superiori al valore 0,50 che rappresenterebbe assenza completa di correlazione o distribuzione casuale, dimostrano correlazioni molto forti e statisticamente significative che richiedono una spiegazione teorica articolata e empiricamente motivata.

Il coefficiente phi (φ), che è una misura standardizzata di associazione per variabili categoriche binarie e che può essere interpretato concettualmente come un coefficiente di correlazione per tabelle di contingenza 2×2, può essere calcolato dalla formula: $\varphi = (n_{11} \times n_{22} - n_{12} \times n_{21}) / \sqrt{[(n_{11} + n_{12})(n_{21} + n_{22})(n_{11} + n_{21})(n_{12} + n_{22})]}$, dove $n_{11} = 420$, $n_{12} = 60$, $n_{21} = 40$, $n_{22} = 380$. Sostituendo questi valori empirici: $\varphi = (420 \times 380 - 60 \times 40) / \sqrt{[(420 + 60)(40 + 380)(420 + 40)(60 + 380)]} = (159.600 - 2.400) / \sqrt{[480 \times 420 \times 460 \times 440]} = 157.200 / \sqrt{[40.161.792.000]} \approx 157.200 / 200.404 \approx 0,784$. Il valore $\varphi \approx 0,78$ indica un'associazione molto forte secondo i criteri standard di interpretazione in statistica (valori di φ superiori a 0,70 sono considerati indicativi di associazioni molto forti in contesti tipologici).

Possiamo anche valutare la significatività statistica di questa associazione mediante il test chi-quadrato (χ^2), che verifica l'ipotesi nulla secondo cui non esisterebbe alcuna associazione sistematica tra le due variabili e le frequenze osservate nelle celle della tabella di contingenza deriverebbero semplicemente da variazione casuale. La statistica chi-quadrato è calcolata come: $\chi^2 = n \times \varphi^2 = 900 \times (0,784)^2 \approx 900 \times 0,615 \approx 553,5$. Con 1 grado di libertà (df = 1 per una tabella 2×2), questo valore di chi-quadrato è estremamente alto e porta a un p-value essenzialmente nullo ($p < 0,0001$), permettendoci di rigettare con confidenza schiacciante l'ipotesi nulla e di concludere che l'associazione osservata è altamente significativa statisticamente e non può essere plausibilmente attribuita a fluttuazioni casuali del campionamento.

Le eccezioni alle correlazioni armoniche, sebbene statisticamente minoritarie costituendo complessivamente solo circa l'11% del campione totale (60 + 40 = 100 lingue su 900), offrono opportunità teoricamente preziose e metodologicamente illuminanti per approfondire ulteriormente la nostra comprensione dei meccanismi causali sottostanti agli universali implicazionali e per identificare fattori modulatori addizionali che possono influenzare le correlazioni tipologiche. Alcune lingue mostrano ordini misti o "disarmonici" che violano apparentemente le previsioni derivanti dagli universali implicazionali nella loro forma più semplice e rigida. L'esistenza documentata e sistematica di queste eccezioni non invalida necessariamente le generalizzazioni tipologiche proposte da Greenberg, ma richiede piuttosto un raffinamento teorico più sofisticato che tenga conto di fattori addizionali e complementari come: primo, l'influenza storica profonda e duratura del contatto linguistico intenso e prolungato con lingue tipologicamente diverse, che può portare al prestito di costruzioni grammaticali e pattern sintattici alieni alla tipologia originaria; secondo, i processi complessi e multifattoriali di cambiamento linguistico diacronico attualmente in corso, dove la lingua si trova in uno stadio transizionale tra due configurazioni tipologiche più stabili; terzo, la stratificazione diacronica di proprietà grammaticali derivanti da stadi evolutivi diversi e successivi della lingua, dove innovazioni recenti coesistono con relitti conservativi di pattern più antichi; quarto, la complessità intrinseca e irriducibile dei sistemi linguistici naturali, che possono manifestare tensioni e competizioni tra principi organizzativi concorrenti e parzialmente incompatibili, portando a compromessi strutturali che violano l'armonia tipologica perfetta.

Il persiano moderno (farsi), la lingua ufficiale dell'Iran contemporaneo parlata da circa 110 milioni di persone come lingua materna o seconda lingua, rappresenta un caso particolarmente interessante, complesso e teoricamente illuminante di lingua che manifesta proprietà tipologiche miste o intermedie che sfidano classificazioni semplicistiche e richiedono analisi attentamente sfumate. Il persiano è tradizionalmente e prevalentemente classificato come lingua SOV per quanto riguarda l'ordine basico non marcato dei costituenti maggiori in frasi transitive dichiarative semplici: "man ketāb rā xāndam" (io libro MARCATORE.OGGETTO lessi = "io lessi il libro"), con ordine lineare superficiale chiaramente e inequivocabilmente soggetto-oggetto-verbo. La particella "rā" (con varianti fonetiche colloquiali come "ro" o "o") marca morfologicamente l'oggetto diretto definito e specifico, disambiguando la funzione sintattica in casi dove l'ordine delle parole potrebbe essere ambiguo o variabile per ragioni pragmatiche o stilistiche.

Tuttavia, diversamente dalla stragrande maggioranza delle lingue SOV prototipiche e armonicamente consistenti come il turco, il giapponese, il coreano, il basco o le lingue dravidiche, il persiano ha sviluppato storicamente e mantiene sincronicamente un sistema significativamente misto di adposizioni che include tanto preposizioni quanto postposizioni, con le preposizioni che sono diventate progressivamente e sistematicamente più dominanti, numerose e produttive nel persiano moderno contemporaneo rispetto agli stadi più antichi e conservativi della lingua documentati in testi del persiano medio (pahlavi) e del persiano classico medievale.

Esempi rappresentativi di preposizioni ampiamente usate nel persiano moderno standard includono: "be" o "beh" (a, verso, indicando direzione o destinazione), "az" (da, di, indicando origine, provenienza o partitività), "barāye" (per, a favore di, indicando beneficiario o scopo), "dar" (in, dentro, indicando locazione interna o inclusione), "bā" (con, indicando compagnia o strumento), "tā" (fino a, indicando limite spaziale o temporale), tutte le quali precedono sistematicamente e rigidamente il loro complemento nominale seguendo il pattern strutturale caratteristico delle preposizioni: "be madrese" (a scuola), "az Tehrān" (da Tehran), "barāye shomā" (per voi/Lei), "dar xāne" (in casa), "bā dust" (con amico), "tā fardā" (fino a domani). Questa presenza quantitativamente significativa e grammaticalmente produttiva di preposizioni in una lingua fondamentalmente e primariamente classificata come SOV sulla base dell'ordine dei costituenti maggiori verbali costituisce una deviazione notevole e teoricamente problematica dall'armonia tipologica pura e stretta prevista dall'Universale 4 di Greenberg, che prederebbe invece la presenza esclusiva o almeno largamente dominante di postposizioni per una lingua con ordine basico SOV.

La spiegazione storico-linguistica più plausibile e empiricamente supportata di questa apparente anomalia tipologica risiede nei processi complessi, multistrato e multifattoriali di cambiamento diacronico che hanno interessato profondamente il persiano nel corso della sua lunga evoluzione documentata che si estende per oltre 2500 anni attraverso diversi stadi evolutivi ben attestati: persiano antico (avestico e persiano achemenide, circa 600-300 a.C.), persiano medio (pahlavi, circa 300 a.C. - 900 d.C.), e persiano moderno (dal 900 d.C. circa fino al presente). Il persiano antico, ricostruito sulla base di iscrizioni monumentali cuneiformi achemenidi e testi religiosi zoroastriani avestici, mostrava un sistema grammaticale significativamente più conservativo e tipologicamente più armonico con l'ordine SOV, includendo un sistema casuale flessivo nominale più ricco (con nominativo, accusativo, genitivo, dativo, ablativo, locativo, strumentale, vocativo), un maggior numero

relativo di costruzioni postposizionali derivate da sostantivi o participi in costruzioni possessive o ablative, e pattern di modificazione nominale più consistentemente prenominali.

Nel corso dei secoli successivi, attraverso complesse dinamiche di evoluzione endogena interna combinata con influenze esogene derivanti dal contatto intenso e prolungato con lingue tipologicamente diverse appartenenti a diverse famiglie genetiche e aree geografiche, il persiano ha subito trasformazioni strutturali profonde e sistematiche. Durante il periodo del persiano medio, coincidente con l'impero sasanide e caratterizzato da intenso contatto con il greco ellenistico, l'aramaico imperiale e successivamente con l'arabo durante e dopo la conquista islamica del VII secolo d.C., il sistema morfologico nominale ha iniziato un processo graduale e progressivo di semplificazione ed erosione, con la perdita graduale delle distinzioni casuali flessive ereditate dall'indoeuropeo e la loro sostituzione funzionale mediante strategie più analitiche basate sull'ordine fisso delle parole e sull'uso di particelle grammaticali e adposizioni.

Il contatto linguistico particolarmente intenso e duraturo con l'arabo, che è diventato la lingua della religione, dell'amministrazione, della scienza e della cultura alta per diversi secoli dopo la conquista islamica, ha portato non solo al massiccio prestito lessicale (si stima che il 40-50% del vocabolario persiano moderno derivi dall'arabo, particolarmente nei domini semantici dell'amministrazione, della religione, della scienza e della cultura astratta), ma anche al prestito e all'adattamento di costruzioni grammaticali e pattern sintattici. L'arabo, essendo una lingua semitica con ordine basico VSO e quindi fondamentalmente VO e head-initial, utilizza sistematicamente e produttivamente preposizioni piuttosto che postposizioni. L'esposizione prolungata a questo modello sintattico attraverso il bilinguismo diffuso, l'educazione formale, la traduzione di testi, e l'incorporazione di espressioni formulaiche arabe nel discorso persiano ha probabilmente facilitato e accelerato lo sviluppo e l'espansione di costruzioni preposizionali nel persiano.

Inoltre, processi endogeni di grammaticalizzazione hanno contribuito significativamente all'emergenza di nuove preposizioni attraverso la rianalisi di sostantivi spaziali, participi verbali e costruzioni seriali. Per esempio, la preposizione "barāye" (per, a favore di) deriva storicamente dalla fusione e grammaticalizzazione della sequenza "barā-ye", composta da "barā" (per, a) combinato con la particella enclitica genitiva "-ye" (che deriva a sua volta dalla particella ezāfe che marca relazioni di modificazione e possesso). Attraverso l'uso frequente in contesti sintattici specifici, questa sequenza originariamente compositiva e analizzabile si è progressivamente opacizzata semanticamente, fonologicamente ridotta, e rianalizzata sincronicamente come una singola unità preposizionale indivisibile.

Nonostante questa espansione quantitativa e funzionale delle preposizioni, il persiano mantiene anche alcune postposizioni, particolarmente quelle che codificano relazioni semantiche più astratte o più specificamente legate alla struttura del discorso. La postposizione "-rā" (varianti colloquiali "-ro", "-o"), già menzionata, marca l'oggetto diretto definito e specifico e segue il sintagma nominale oggetto: "ketāb-rā" (libro-ACC.DEF = "il libro" in funzione di oggetto). Altre postposizioni includono forme derivanti da sostantivi direzionali grammaticalizzati che possono ancora mostrare una certa variabilità posizionale.

Per quanto riguarda la posizione dei modificatori genitivi, il persiano manifesta ordini variabili e flessibili che dipendono da fattori lessicali, semantici, sintattici e stilistici complessi, riflettendo ancora una volta la natura tipologicamente mista della lingua. La costruzione genitiva più produttiva e frequente utilizza la particella enclitica "e" (ezāfe, con varianti "ye" dopo vocali), che si aggiunge al nome testa e lega sintatticamente e prosodicamente il modificatore seguente. Questa costruzione tipicamente segue l'ordine Nome-EZAFE-Modificatore, con il modificatore genitivo in posizione postnominale: "ketāb-e Ali" (libro-EZAFE Ali = "il libro di Ali"), "xāne-ye man" (casa-EZAFE io = "la mia casa"), "doxtar-e ostād" (figlia-EZAFE professore = "la figlia del professore").

Questo ordine Nome-Genitivo, con genitivo postnominale, è armonico con la presenza dominante di preposizioni nella lingua contemporanea, confermando l'Universale 2 nella sua forma condizionale: se preposizioni, allora genitivo postnominale. Tuttavia, il persiano permette anche costruzioni possessive alternative con ordine invertito in contesti stilisticamente marcati o arcaizzanti, particolarmente nella poesia classica e in espressioni idiomatiche fossilizzate, dove il possessore può precedere il nome posseduto senza la particella ezāfe, seguendo un pattern più antico e presumibilmente più originario: "Ali ketāb" (Ali libro = "il libro di Ali", ordine marcato e arcaico). Questa variazione e questa stratificazione di pattern riflettono la transizione storica del persiano da un sistema più armonicamente OV verso un sistema progressivamente più misto.

Inoltre, il persiano mostra variabilità notevole nella posizione di altri modificatori nominali oltre ai genitivi. Gli aggettivi qualificativi mostrano forte preferenza per la posizione postnominale nel persiano moderno standard: "ketāb-e qermez" (libro-EZAFE rosso = "libro rosso"), "xāne-ye bozorg" (casa-EZAFE grande = "casa grande"), ancora una volta utilizzando la particella ezāfe per legare il modificatore aggettivale alla testa nominale. Questo ordine Nome-Aggettivo è armonico con le preposizioni e con il genitivo postnominale, suggerendo una deriva generale verso pattern più consistentemente head-initial nel dominio nominale.

Tuttavia, alcune classi ristrette di aggettivi, particolarmente quelli con funzione più determinativa o quantificazionale che qualificativa, possono precedere il nome, mostrando ancora tracce del pattern più antico head-final: "har ketāb" (ogni libro), "tamām-e ketābhā" (tutto-EZAFE libri = "tutti i libri"). I numerali, criticamente, mostrano posizione categoricamente prenominale: "se ketāb" (tre libro = "tre libri"), "dah dānešju" (dieci studente = "dieci studenti"), "sad sāl" (cento anno = "cento anni"), conformandosi all'Universale 18 di Greenberg che prevede che se aggettivi sono prenominali, allora numerali sono necessariamente prenominali, ma permettendo anche la configurazione inversa dove numerali possono essere prenominali anche quando aggettivi sono prevalentemente postnominali, riflettendo la posizione gerarchica più nucleare dei numerali nella struttura del sintagma nominale.

Le clausole relative nel persiano moderno seguono rigidamente e categoricamente il nome testa, manifestando ordine Nome-[Clausola Relativa] in perfetta armonia con le preposizioni dominanti e l'ordine generale progressivamente più head-initial nel dominio nominale: "ketābi ke xāndam" (libro che lessi = "il libro che lessi"), "mardi ke diruz didām" (uomo che ieri vidi = "l'uomo che vidi ieri"), "xāneyi ke dar ān zendegi mikonam" (casa che in essa vivere fare.PRESENTE-1SG = "la casa nella quale vivo"). Le clausole relative sono

introdotte dal complementatore "ke" (che) e seguono la testa nominale, che può essere marcata dal suffisso indefinito "-i" quando la testa è indefinita. La struttura interna delle clausole relative mantiene l'ordine SOV caratteristico del persiano: dopo il complementatore, gli argomenti nominali precedono il verbo ("diruz didām" = ieri vidi, con ordine Avverbio-Oggetto.implicito-Verbo).

Questa configurazione apparentemente contraddittoria e tipologicamente mista del persiano, dove l'ordine basico SOV nelle frasi principali coesiste con preposizioni dominanti, genitivi e aggettivi prevalentemente postnominali, e clausole relative postnominali, ha generato un dibattito teorico considerevole, prolungato e metodologicamente sofisticato nella letteratura tipologica e sintattica. Alcune analisi hanno proposto che il persiano debba essere rianalizzato come lingua in transizione attiva da un tipo più puramente SOV e armonicamente head-final, preservato negli stadi più antichi della lingua, verso un tipo progressivamente più misto o addirittura tendenzialmente SVO con caratteristiche head-initial espanse, riflettendo influenze storiche cumulative dall'arabo, dal greco, dal turco, e dalle dinamiche interne di grammaticalizzazione e rianalisi.

Altri studiosi hanno osservato che la classificazione basica SOV versus SVO potrebbe essere inadeguata o troppo rigida per catturare la complessità delle lingue con ordini multipli e flessibilità pragmatica. Il persiano moderno colloquiale mostra una certa variabilità nell'ordine dei costituenti maggiori, con ordine SVO che diventa progressivamente più frequente in registri informali e nel parlato spontaneo, particolarmente quando l'oggetto è indefinito e non marcato dalla particella "rā": "man ketāb xāndam" (io libro lessi, SOV) versus "man xāndam ketāb" (io lessi libro, SVO), entrambi accettabili con sottili differenze pragmatiche di topicalizzazione e focus informazionale.

Questa flessibilità e questa variazione suggeriscono che il persiano possa essere concettualizzato non come una lingua con un singolo ordine basico rigido e inviolabile, ma piuttosto come una lingua con ordini multipli condizionati da fattori pragmatici, informazionali e stilistici, dove diverse configurazioni sintattiche coesistono nel repertorio grammaticale della comunità linguistica e vengono selezionate contestualmente secondo principi complessi di struttura informazionale. Da questa prospettiva, le apparenti anomalie tipologiche del persiano riflettono non tanto violazioni degli universali implicazionali quanto la complessità e la multidimensionalità dei sistemi linguistici reali, che possono manifestare simultaneamente proprietà associabili a diversi tipi ideali nella classificazione tipologica.

Le Clausole Relative e la Gerarchia di Accessibilità: Un'Analisi Approfondita dell'Universale 24 e delle Sue Ramificazioni Teoriche Complesse

Le clausole relative rappresentano indiscutibilmente uno dei terreni di analisi più ricchi, complessi, strutturalmente sofisticati e teoricamente illuminanti per lo studio sistematico degli universali linguistici e delle correlazioni tipologiche cross-linguistiche. Una clausola relativa è una clausola subordinata o dipendente che modifica un sintagma nominale (la cosiddetta "testa" della costruzione relativa, terminologicamente distinta dal modificatore stesso) fornendo informazioni aggiuntive, qualificative, restrittive o non restrittive, su di esso.

Le clausole relative sono strutture sintatticamente complesse e articolate che coinvolgono tipicamente un elemento grammaticale relativizzatore (che può essere un pronome relativo declinabile secondo caso, genere e numero come nelle lingue indoeuropee classiche, un complementatore invariabile come in molte lingue moderne, o anche nessuna marcatura esplicita in alcune lingue con strategie di relativizzazione mediante giustapposizione o gap) e una lacuna o gap strutturale corrispondente a un argomento, complemento o aggiunto sintattico mancante nella clausola subordinata, il cui referente viene identificato semanticamente e co-indicizzato sintatticamente con la testa nominale esterna alla clausola relativa.

Per fornire un'illustrazione concreta e analiticamente dettagliata da una lingua romanza prototipica, consideriamo la frase italiana complessa "il libro che ho letto ieri sera", che contiene una clausola relativa restrittiva ("che ho letto ieri sera") che modifica e restringe semanticamente il riferimento del sintagma nominale testa "il libro". La struttura interna della clausola relativa può essere analizzata sintatticamente come contenente un complementatore invariabile "che" (che non manifesta flessione per caso, genere o numero, a differenza dei pronomi relativi pienamente declinabili del latino classico come "qui, quae, quod") e una struttura clausale subordinata con una posizione argomentale oggetto sintattica vuota o gap strutturale: "[che [io ho letto ___ ieri sera]]", dove il gap (___) corrisponde alla posizione canonica dell'oggetto diretto del verbo transitivo "letto" ed è semanticamente co-indicizzato e sintatticamente legato alla testa nominale esterna "libro" attraverso la mediazione del complementatore relativo.

La posizione lineare superficiale delle clausole relative rispetto alla testa nominale che modificano mostra una variazione tipologica significativa, sistematica e strutturatamente correlata con altri aspetti indipendenti dell'ordine delle parole nelle lingue del mondo, e, cosa teoricamente cruciale e metodologicamente centrale per la nostra discussione approfondita, questa posizione correla robustatamente e predicibilmente con altri aspetti dell'ordine dei costituenti secondo pattern implicazionali identificati originariamente da Greenberg nel suo lavoro seminale del 1963 e successivamente raffinati, estesi, quantificati e teoricamente elaborati nella letteratura tipologica e sintattica formale delle decadi successive.

L'Universale 24, nella sua formulazione originale di Greenberg che ha subito varie interpretazioni concorrenti, raffinamenti metodologici e articolazioni teoriche alternative nella letteratura successiva, afferma nella sua forma più diretta e letterale che se la posizione relativa di soggetto e oggetto nella frase transitiva basica è SOV (ovvero se l'oggetto diretto precede linearmente il verbo nella serializzazione superficiale canonica), allora la lingua ha sempre o quasi sempre clausole relative prenominali, cioè clausole relative che precedono sistematicamente e categoricamente il nome testa che modificano, manifestando l'ordine lineare [Clausola Relativa] Nome piuttosto che l'ordine inverso e speculare Nome [Clausola Relativa]. In altre parole, e formulato in termini di correlazione tipologica piuttosto che di affermazione assoluta, le lingue in cui l'oggetto precede il verbo nella frase principale transitiva tendono sistematicamente, robustatamente e con alta probabilità condizionale a collocare le clausole relative subordinate modificatrici prima del nome testa che modificano, generando una configurazione strutturale dove il modificatore clausale complesso precede la testa nominale modificata.

Formalmente, possiamo esprimere questo universale implicazionale come una probabilità condizionale quantitativa: P(Clausole_relative_prenominali | SOV) ≈ 0,85-0,90, dove la notazione P(A | B) denota come convenzione standard la probabilità dell'evento A condizionata al verificarsi dell'evento B. Questo valore probabilistico empiricamente stimato indica che, data una lingua con ordine basico SOV verificato attraverso l'analisi di frasi transitive semplici non marcate, la probabilità che quella stessa lingua abbia clausole relative prenominali è molto alta, dell'ordine dell'85-90% secondo le stime più affidabili, metodologicamente controllate e empiricamente robuste basate su campioni tipologici ampi, geneticamente diversificati e geograficamente distribuiti. La deviazione sostanziale e statisticamente significativa da 0,50 (che rappresenterebbe assenza completa di correlazione o distribuzione puramente casuale e indipendente) dimostra una forte associazione statistica positiva che richiede esplicazione teorica.

Questo pattern generalizzato si osserva con notevole regolarità, consistenza cross-linguistica impressionante e robustezza empirica in numerose lingue prototipicamente e inequivocabilmente SOV appartenenti a diverse e distinte famiglie genetiche ampiamente separate filogeneticamente e distribuite in diverse aree geografiche distanti e non contigue, minimizzando quindi l'ipotesi di diffusione areale per contatto come spiegazione primaria della correlazione. Le famiglie e lingue rilevanti includono: lingue altaiche come turco, azero, turkmeno, uzbeko, kazako, kirghiso, tataro (ramo turco), mongolo, buriato, calmucco (ramo mongolico), evenki, manciù, nanai (ramo tunguso); lingue dravidiche meridionali come tamil, telugu, kannada, malayalam; lingue sino-tibetane come birmano, tibetano, lepcha, newari; lingue uralico-finniche come finlandese, estone, careliano, vepso, ingermanlandese, livone (ramo finnico), ungherese (ramo ugrico); lingue isolate geneticamente non classificabili come giapponese, coreano, basco, burushaski; lingue indo-iraniane orientali e centrali come hindi, urdu, punjabi, gujarati, bengali, marathi, nepali; e molte altre lingue geneticamente diverse ma convergenti tipologicamente verso il pattern SOV con clausole relative prenominali.

Consideriamo un'analisi strutturale dettagliata, morfologicamente trasparente e teoricamente illuminante con esempi linguistici specifici e accuratamente glosati dal turco ottomano e moderno, una lingua altaica meridionale SOV prototipica e armonicamente consistente con clausole relative prenominali particolarmente chiare, regolari e morfologicamente esplicite nella loro formazione. La frase italiana "il libro che ho letto" corrisponde in turco moderno standard contemporaneo alla forma nominalizzata compatta "okuduğum kitap", che può essere analizzata e segmentata morfema per morfema secondo la struttura agglutinativa caratteristica delle lingue altaiche come: "oku-duğ-um kitap". Analizziamo sistematicamente ogni componente morfologico: "oku-" è la radice lessicale verbale non flessiva "leggere" (infinito okumak); "-duğ-" (con varianti allomorfe "-dık/-dik/-duk/-dük" secondo le regole sistematiche e predicibili di armonia vocalica palatale/velare e arrotondata/non-arrotondata che governano la fonologia turca) è un suffisso derivazionale di nominalizzazione o participializzazione che converte la radice verbale in un nome verbale astratto (deverbal noun) o participio relativamente (relative participle); "-um" è un suffisso possessivo personale flessivo di prima persona singolare (varianti -ım/-im/-um/-üm secondo armonia vocalica) che marca morfologicamente l'agente implicito o soggetto logico della nominalizzazione, indicando che "io" sono colui che ha eseguito l'azione di leggere; "kitap" (con pronuncia [kitap] o [kitab] in varianti dialettali) significa "libro" e costituisce la testa nominale esterna modificata dalla clausola relativa nominalizzata precedente.

La struttura morfosintattica completa può essere rappresentata gerarchicamente come: [[oku-duğ-um]Clausola.Relativa.Nominalizzata [kitap]Testa.Nominale]Sintagma.Nominale. La clausola relativa nominalizzata "okuduğum" precede chiaramente, categoricamente e invariabilmente la testa nominale "kitap", manifestando l'ordine prenominale sistematico caratteristico delle lingue SOV armoniche. Non esiste possibilità grammaticale in turco standard di invertire questo ordine (*"kitap okuduğum" sarebbe completamente agrammaticale e incomprensibile come costruzione relativa, sebbene possa essere interpretato come sequenza di due sintagmi nominali giustapposti in contesti particolari).

La struttura interna della clausola relativa nominalizzata riflette perfettamente e preserva fedelmente l'ordine SOV caratteristico e dominante del turco nelle frasi principali finite. In una frase principale transitiva indipendente corrispondente come "io ho letto il libro" o "io lessi il libro", il turco usa invariabilmente l'ordine SOV: "(ben) kitabı okudum", che può essere segmentato come: "ben" (pronome personale di prima persona singolare nominativo "io", spesso omesso per pro-drop dato che la flessione verbale marca univocamente la persona) + "kitab-ı" (libro-ACCUSATIVO, dove il suffisso casuale accusativo -ı/-i/-u/-ü marca morfologicamente ed esplicitamente l'oggetto diretto definito) + "oku-du-m" (leggere-PASSATO-1SG, dove -du/-dü è il suffisso di tempo passato definitivo e -m è il suffisso di accordo personale per la prima persona singolare). L'ordine risultante è chiaramente Soggetto(implicito)-Oggetto-Verbo.

Quando questa struttura transitiva basica SOV viene nominalizzata e trasformata morfosintatticamente per funzionare come modificatore di una testa nominale esterna creando una clausola relativa, la configurazione interna SOV viene preservata e mantenuta intatta, semplicemente incorporando e subordinando l'intera struttura verbale come modificatore prenominale della testa attraverso il processo morfologico produttivo e altamente regolare di nominalizzazione participiale. Se l'oggetto è esplicitamente presente all'interno della clausola relativa (quando la relativizzazione avviene su una posizione diversa dall'oggetto, ad esempio sul soggetto o su un obliquo), esso mantiene la sua posizione prenominale canonica: "kitabı okuduğum adam" (libro-ACCUSATIVO leggere-NOMINALIZZATORE-POSSESSIVO.1SG uomo = "l'uomo che lesse il libro", dove la relativizzazione è sul soggetto "uomo" e l'oggetto "libro" appare esplicitamente all'interno della relativa mantenendo ordine OV).

Consideriamo un esempio strutturalmente più complesso, sintatticamente articolato e teoricamente più illuminante per illustrare ulteriormente la robustezza e la sistematicità del pattern. La frase italiana elaborata "il libro che l'uomo ha dato alla donna ieri" si rende in turco come la costruzione nominalizzata complessa "adamın kadına dün verdiği kitap", che può essere analizzata morfema per morfema e costituente per costituente come: "adam-ın" (uomo-GENITIVO, dove il suffisso casuale genitivo -ın/-in/-un/-ün marca il possessore o agente della nominalizzazione), "kadın-a" (donna-DATIVO, dove il suffisso casuale dativo -a/-e marca il recipiente o beneficiario), "dün" (avverbio temporale "ieri"), "ver-diğ-i" (dare-NOMINALIZZATORE-POSSESSIVO.3SG, dove il suffisso possessivo di terza persona singolare -i/-ı/-u/-ü marca l'accordo con l'agente genitivo "adamın"), "kitap" (libro, testa nominale esterna).

Tutta la clausola relativa complessa e sintatticamente articolata "adamın kadına dün verdiği", che contiene tre argomenti o aggiunti (agente, recipiente, tempo) più il verbo nominalizzato,

precede rigidamente e categoricamente il nome testa "kitap", mantenendo internamente la configurazione canonica SOV del turco: il soggetto/agente "adamın" precede l'oggetto indiretto/recipiente "kadına" che precede l'avverbio temporale "dün" che precede il verbo nominalizzato "verdiği". Questa consistenza strutturale multi-livello, che si manifesta sia nell'ordine esterno della clausola relativa rispetto alla testa (prenominale) sia nell'ordine interno dei costituenti all'interno della clausola relativa (SOV), dimostra in modo particolarmente chiaro e impressionante l'armonia tipologica profonda, sistematica e pervasiva della lingua turca attraverso diversi livelli e domini della struttura grammaticale.

Il giapponese, un'altra lingua SOV prototipica, archetipica e armonicamente consistente geneticamente isolata e non correlata filogeneticamente con il turco, mostra pattern strutturalmente strettamente paralleli e tipologicamente convergenti, fornendo evidenza indipendente e cross-linguisticamente robusta per la validità dell'universale. La frase italiana "il libro che ho letto" si rende in giapponese standard contemporaneo come "watashi ga yonda hon", che può essere analizzata come: "watashi" (pronome personale di prima persona "io"), "ga" (particella grammaticale invariabile che marca il soggetto/nominativo), "yon-da" (leggere-PASSATO, dove "yomu" è l'infinito e "yonda" è la forma passata irregolare per contrazione fonetica), "hon" (sostantivo "libro", testa nominale esterna).

La clausola relativa "watashi ga yonda" (io NOMINATIVO leggere-PASSATO = "che io lessi") precede categoricamente e invariabilmente la testa "hon" (libro), manifestando ordine prenominale. Notare che il giapponese, a differenza del turco che usa nominalizzazione esplicita attraverso morfologia derivazionale complessa con affissi specializzati, usa semplicemente una forma verbale finita standard (il passato "yonda") senza marcatura morfologica speciale, affissi relativizzatori dedicati o complementatori espliciti per introdurre e subordinare la clausola relativa, riflettendo una strategia di relativizzazione tipologicamente diversa basata su giustapposizione diretta e gap strutturale piuttosto che su nominalizzazione esplicita, ma mantenendo comunque la stessa posizione lineare prenominale in perfetta armonia con l'ordine generale OV della lingua.

Il pattern opposto, inverso e speculare si osserva sistematicamente, robustatamente e predicibilmente nelle lingue VO, in particolare quelle con ordine basico SVO come l'italiano, l'inglese, il francese, lo spagnolo, il portoghese, il rumeno e la maggior parte delle altre lingue romanze e germaniche moderne, così come molte lingue slave, baltiche e di altre famiglie. In queste lingue tipologicamente VO e tendenzialmente head-initial, le clausole relative seguono tipicamente, preferenzialmente e nella stragrande maggioranza dei casi il nome testa che modificano, manifestando l'ordine lineare superficiale Nome [Clausola Relativa] in armonia strutturale e direzionale con l'ordine generale head-initial o testa-complemento che caratterizza queste lingue sia nel dominio verbale (V-O order) sia in molteplici altri domini sintattici indipendenti. Formalmente, questa correlazione complementare e speculare può essere espressa come: P(Clausole_relative_postnominali | VO) ≈ 0,75-0,85, indicando una correlazione forte anche se leggermente meno robusta e predicibile di quella osservata nelle lingue OV, riflettendo la maggiore variabilità tipologica, flessibilità strutturale e tendenza a configurazioni miste che caratterizza le lingue VO (e specialmente il sottotipo SVO) rispetto alle lingue OV più armonicamente coerenti.

In italiano standard contemporaneo, come già osservato e illustrato precedentemente, le clausole relative sono sistematicamente, categoricamente e senza eccezioni significative

postnominali in tutte le varietà diastratiche, diatopiche e diafasiche: "il libro [che ho letto]", "la persona [che è arrivata]", "lo studente [a cui ho parlato]", "il paese [da cui vengo]", "il giorno [in cui ci siamo incontrati]", dove le clausole relative subordinate racchiuse in parentesi quadre per chiarezza espositiva seguono sempre e invariabilmente il sintagma nominale testa, che può includere l'articolo determinativo o indeterminativo, aggettivi prenominali o postnominali, e altri modificatori secondo la struttura complessa del sintagma nominale esteso italiano.

La struttura interna sintattica delle clausole relative italiane riflette e preserva l'ordine SVO caratteristico e basico dell'italiano nelle frasi principali indipendenti: dopo il pronome relativo morfologicamente variabile ("che" invariabile per soggetto/oggetto, "cui" per obliqui preceduti da preposizione, "il quale/la quale/i quali/le quali" come forme alternative più formali e flessionalmente esplicite) o il complementatore invariabile "che" che introduce la clausola subordinata, segue una struttura clausale con ordine soggetto-verbo(-oggetto) quando il gap strutturale non occupa la posizione soggetto. Per esempio, in "il libro [che io ho letto]", la struttura interna dopo "che" è "io ho letto ___" con ordine Soggetto-Verbo-Gap(Oggetto), preservando la sequenza SV caratteristica. In "il libro [che mi piace]", la struttura è "che ___ mi piace" con gap in posizione soggetto seguito da verbo, ancora preservando l'ordine basico della lingua.

L'inglese mostra pattern strutturalmente paralleli e tipologicamente convergenti con clausole relative sistematicamente postnominali: "the book [that I read]", "the book [which I read]" (più formale), "the book [I read]" (con omissione del relativizzatore when the gap is object), "the person [who arrived]", "the student [to whom I spoke]" (formale), "the student [I spoke to]" (informale con preposition stranding). La struttura interna mantiene rigorosamente l'ordine SVO basico dell'inglese: "that I read ___" ha ordine soggetto-verbo con gap in posizione oggetto posverbale, mentre "who ___ arrived" ha gap in posizione soggetto preverbal seguito dal verbo, mantenendo sempre la posizione relativa V rispetto agli argomenti nominali quando presenti.

Il francese standard: "le livre [que j'ai lu]" (il libro che io-ho letto), "la personne [qui est arrivée]" (la persona che è arrivata), "l'étudiant [à qui j'ai parlé]" (lo studente a cui io-ho parlato). Lo spagnolo: "el libro [que leí]" (il libro che lessi), "la persona [que llegó]" (la persona che arrivò), "el estudiante [a quien hablé]" (lo studente a chi parlai). Il portoghese: "o livro [que li]", "a pessoa [que chegou]", "o estudante [a quem falei]". Il rumeno: "cartea [pe care am citit-o]" (il-libro PE che ho letto-lo, con marcatura accusativa speciale "pe" e clitic resumptive), "persoana [care a sosit]", "studentul [cu care am vorbit]".

In tutti questi casi rappresentativi di lingue romanze geneticamente correlate ma indipendentemente attestanti lo stesso pattern, la clausola relativa segue invariabilmente e categoricamente il sintagma nominale testa in perfetta armonia strutturale con l'ordine generale VO/SVO della lingua, e la struttura interna clausale mantiene l'ordine VO/SVO caratteristico, manifestando consistenza direzionale multi-livello head-initial pervasiva attraverso diversi domini grammaticali.

Questa correlazione robusta, sistematica e cross-linguisticamente attestata tra ordine OV/VO nella frase principale transitiva e posizione delle clausole relative (prenominale versus postnominale) non è ovviamente assoluta, deterministica e senza eccezioni nel

senso più rigido, ma piuttosto statistica, probabilistica e tendenziale come già enfatizzato ripetutamente, e le eccezioni apparenti, i casi problematici, le configurazioni miste e le lingue tipologicamente intermedie offrono opportunità particolarmente preziose, teoricamente illuminanti e metodologicamente fruttuose per raffinare sostanzialmente e approfondire significativamente la nostra comprensione teorica dei fattori causali multipli, delle pressioni strutturali concorrenti e dei principi organizzativi complessi che sono in gioco nella determinazione dei pattern tipologici osservati.

Una complicazione teorica importante, empiricamente significativa e ampiamente discussa nella letteratura tipologica e sintattica formale degli ultimi quattro decenni deriva dal fatto che alcune lingue, come il cinese mandarino menzionato e analizzato precedentemente, sono classificate come SVO per quanto riguarda l'ordine superficiale dei costituenti maggiori della frase transitiva basica non marcata sulla base di criteri diagnostici standard, ma mostrano clausole relative prenominali in apparente violazione o controesempio all'Universale 24 di Greenberg nella sua formulazione più diretta e letterale.

La frase italiana "il libro che ho letto" si rende in cinese mandarino standard contemporaneo (pǔtōnghuà) come la sequenza compatta "wǒ dú de shū", che può essere analizzata morfema per morfema e costituente per costituente come: "wǒ" (pronome personale di prima persona singolare "io"), "dú" (verbo lessicale "leggere" nella sua forma basica non flessa, dato che il cinese è tipologicamente una lingua isolante senza morfologia flessiva verbale ricca), "de" (particella grammaticale multifunzionale di modificazione nominalizzante, relativizzante o subordinante), "shū" (sostantivo "libro", testa nominale esterna). Analizziamo la struttura sintattica: "wǒ" è il soggetto o agente della clausola relativa; "dú" è il verbo transitivo; "de" è una particella funzionale grammaticalizzata che marca la relazione di modificazione tra la clausola precedente e il nome seguente; "shū" è la testa nominale modificata.

La clausola relativa "wǒ dú" (io leggere = "che io lessi/leggo") precede chiaramente e categoricamente la testa nominale "shū" (libro), manifestando ordine prenominale strutturalmente identico a quello delle lingue SOV prototipiche come turco e giapponese, nonostante il fatto che il cinese mandarino sia consensualmente classificato come lingua SVO sulla base dell'ordine dei costituenti maggiori nelle frasi principali finite e indipendenti.

Questo apparente controesempio significativo ha generato, alimentato e perpetuato un dibattito teorico considerevole, prolungato e metodologicamente sofisticato nella letteratura tipologica e sintattica formale delle ultime cinque decadi. Diverse analisi concorrenti, proposte esplicative alternative e rianalizzazioni strutturali sono state avanzate, proposte e difese da diversi studiosi e scuole teoriche per spiegare, riconciliare o rianalizzare questo caso empiricamente problematico e teoricamente cruciale.

Una prima linea di analisi, sviluppata particolarmente nei lavori influenti di Charles Li, Sandra Thompson, James Huang e altri sinologi-linguisti, ha proposto che il cinese mandarino, nonostante l'ordine SVO superficiale predominante nelle frasi principali monoclausal semplici con verbi transitivi finiti, mostri caratteristiche strutturali profonde, tipologiche sottostanti e pattern sintattici multipli di tipo OV in molti e diversi domini indipendenti della grammatica oltre alle clausole relative, suggerendo che la classificazione simplicistica come

lingua puramente SVO possa essere inadeguata, superficiale o metodologicamente fuorviante.

Le proprietà tipologicamente OV del cinese mandarino includono, oltre alle clausole relative prenominali già discusse: primo, la presenza e l'uso produttivo di postposizioni coesistenti con preposizioni, creando un sistema misto di adposizioni dove entrambi i tipi sono attestati, produttivi e funzionalmente specializzati, sebbene le preposizioni siano numericamente più comuni e categorialmente più produttive; secondo, la tendenza generale, sistematica e pervasiva a collocare modificatori sintattici prima delle teste sintattiche nel dominio nominale esteso, inclusi aggettivi qualificativi che precedono categoricamente il nome ("hóng shū" = rosso libro = "libro rosso", mai *"shū hóng"), genitivi e sintagmi possessivi che precedono il nome ("Zhāngsān de shū" = Zhangsan PARTICELLA libro = "il libro di Zhangsan"), numerali con classificatori che precedono il nome ("sān běn shū" = tre CLASSIFICATORE libro = "tre libri"), dimostrativi che precedono il nome ("zhè běn shū" = questo CLASSIFICATORE libro = "questo libro"), e clausole relative che precedono il nome come già discusso; terzo, la presenza di costruzioni verbali seriali e risultative dove il verbo principale precede il verbo sussidiario o risultativo in pattern che assomigliano più a configurazioni OV che VO in certe analisi strutturali.

Secondo questa prospettiva analitica influente e ampiamente discussa, il cinese mandarino dovrebbe essere concettualizzato non come una lingua SVO pura e prototipica del tipo romanzo-germanico occidentale, ma piuttosto come una lingua con ordine misto, transizionale o ibrido che combina proprietà SVO nel dominio della frase principale (ordine dei costituenti maggiori verbali) con proprietà OV persistenti e sistematiche nel dominio nominale e in altri domini strutturali, possibilmente riflettendo un'evoluzione storica complessa e stratificata.

Evidenza diacronica comparativa dal cinese arcaico (Old Chinese, utilizzato nei testi classici del periodo pre-Qin anteriore al 221 a.C.) e dal cinese classico (Classical Chinese o Literary Chinese, utilizzato nei testi scritti fino al ventesimo secolo) suggerisce, secondo ricostruzioni filologiche e analisi sintattiche di testi antichi, che stadi più antichi della lingua cinese potrebbero aver manifestato ordini diversi, più flessibili o variabili, con una possibile maggiore prominenza relativa di pattern SOV o almeno ordini misti in certi contesti sintattici, registri stilistici o costruzioni particolari.

Alcuni studiosi hanno osservato che il cinese arcaico potrebbe aver avuto maggiore flessibilità nell'ordine dei costituenti con una certa frequenza non trascurabile di ordini SOV o OSV in contesti pragmaticamente marcati o poetici, e che l'ordine SVO relativamente rigido del mandarino moderno standard potrebbe rappresentare un'innovazione relativamente recente (sviluppatasi progressivamente nel corso del primo millennio d.C.) piuttosto che la configurazione originaria prototipica ereditata dal sino-tibetano comune ricostruito. Da questa prospettiva storico-comparativa, le clausole relative prenominali del cinese mandarino moderno potrebbero essere interpretate come relitti conservativi o persistenze strutturali di un pattern tipologico OV più antico e originario, preservato specificamente nel dominio della modificazione nominale clausale anche dopo che l'ordine generale della frase principale si è spostato progressivamente verso SVO attraverso processi diacronici di cambiamento sintattico graduale.

Una seconda linea di analisi, sviluppata particolarmente in lavori di sintassi formale generativista e cartografica, ha osservato che le cosiddette "clausole relative" del cinese mandarino sono strutturalmente, categorialmente e morfosintatticamente diverse e non completamente parallele a quelle delle tipiche lingue SVO indoeuropee come italiano, inglese o francese in modi teoricamente significativi e analiticamente cruciali. Piuttosto che essere vere clausole subordinate completamente verbali con struttura clausale completa, articolata e massimalmente proiettata contenente un pronome relativo o complementatore che introduce una frase finita con tempo, aspetto, mood e accordo pienamente specificati, le costruzioni relative del cinese mandarino sono morfologicamente, sintatticamente e semanticamente più simili a strutture di nominalizzazione compatte, costruzioni partecipiali ridotte o modificatori aggettivali derivati, avvicinandosi categorialmente e funzionalmente più agli aggettivi modificatori semplici o ai participi nominalizzati che alle clausole subordinate finite completamente verbali e sintatticamente autonome.

La particella "de" (的) che marca e introduce queste costruzioni modificatrici nel cinese mandarino è estremamente multifunzionale, grammaticalizzata e semanticamente generale, apparendo in una vastissima gamma di contesti di modificazione nominale, possessione, relativizzazione e nominalizzazione, non essendo affatto limitata o specializzata alle sole clausole relative nel senso stretto e tecnico del termine. La particella "de" appare: nelle costruzioni genitive e possessive ("wǒ de shū" = io PARTICELLA libro = "il mio libro"), nelle costruzioni di modificazione aggettivale ("hóng de shū" = rosso PARTICELLA libro = "libro rosso", quando l'aggettivo è enfatizzato o contrastato), nelle costruzioni di nominalizzazione ("chī de" = mangiare PARTICELLA = "cosa da mangiare" o "cibo"), nelle costruzioni di complemento risultativo nominalizzato, e naturalmente nelle costruzioni relativizzanti ("wǒ dú de shū" = io leggere PARTICELLA libro = "libro che io leggo/lessi").

Questa distribuzione estremamente ampia, multifunzionale e categorialmente diversificata della particella "de" supporta l'idea che essa sia fondamentalmente e primariamente un marcatore generale di modificazione e relazione tra due elementi nominali o tra un elemento predicativo e un nome, piuttosto che un complementatore relativizzante specializzato e funzionalmente dedicato alla subordinazione di clausole relative finite nel senso tecnico-sintattico stretto. La struttura che precede "de" nelle costruzioni relativizzanti può essere analizzata plausibilmente non come una clausola verbale finita completamente proiettata con tutti i livelli funzionali della periferia sinistra (CP, TP, AspP, AgrP), ma piuttosto come una proiezione verbale ridotta, una struttura partecipiale o nominalizzata, o addirittura semplicemente come una sequenza verbale minimale senza proiezioni funzionali estese.

Questa differenza strutturale profonda, categoriale e architetturale potrebbe spiegare plausibilmente perché le costruzioni relativizzanti del cinese seguono principi di collocazione posizionale diversi e indipendenti rispetto alle vere clausole relative finite e completamente verbali delle lingue indoeuropee: se sono strutturalmente e categorialmente più vicine ai modificatori aggettivali nominalizzati o ai participi derivazionali che alle clausole subordinate complete, la loro posizione prenominale sarebbe perfettamente consistente con la posizione prenominale generale di tutti i modificatori nominali nel cinese (aggettivi, genitivi, numerali, dimostrativi), riflettendo semplicemente il principio più generale che tutti i modificatori nel dominio nominale del cinese precedono la testa, indipendentemente dal fatto che la lingua sia SVO nel dominio verbale.

Una terza linea di analisi, sviluppata in lavori di sintassi formale minimalista e derivazionale, ha proposto che le costruzioni relativizzanti prenominali del cinese non siano derivate sintatticamente per movimento di un operatore relativo nullo o visibile da una posizione interna alla clausola verso la periferia sinistra con creazione di una dipendenza operator-variable mediante movimento-A-barra, come nelle tipiche relative delle lingue indoeuropee analizzate secondo il quadro teorico chomskiano standard, ma siano piuttosto semplici clausole modificatrici aggiunte in posizione di specificatore o adjoined direttamente al sintagma nominale senza la struttura operatore-variabile caratteristica e definitoria delle vere relative con movimento.

Secondo queste analisi formali alternative, la differenza strutturale profonda nella derivazione sintattica (modificatore aggiunto versus clausola con operatore mosso) predicherebbe naturalmente differenze nella posizione lineare superficiale: le costruzioni con operatore mosso tipicamente seguono la testa nelle lingue VO perché l'operatore si muove verso la periferia sinistra della clausola subordinata che è a sua volta collocata a destra della testa per armonia direzionale, mentre le costruzioni con modificatore aggiunto possono precedere la testa seguendo semplicemente i principi generali di aggiunta dei modificatori nella lingua, che nel cinese sono sistematicamente prenominali.

Questa molteplicità di analisi teoriche concorrenti e proposte esplicative alternative riflette la complessità empirica profonda e la ricchezza teorica del caso del cinese mandarino, che continua a essere un terreno di dibattito attivo e metodologicamente produttivo nella letteratura tipologica e sintattica contemporanea. Piuttosto che costituire semplicemente un controesempio problematico che falsifica l'Universale 24, il cinese può essere visto più costruttivamente come un caso che rivela limitazioni delle classificazioni tipologiche troppo semplicistiche e rigide, che dimostra l'importanza di distinguere attentamente tra diversi tipi strutturali di costruzioni modificatrici superficialmente simili, e che sottolinea la necessità di analisi fine-grained e strutturalmente informate piuttosto che di generalizzazioni puramente basate su ordini lineari superficiali.

I Numerali e i Nomi: L'Universale 18 e la Gerarchia Implicazionale dei Modificatori Nominali

Un altro dominio empiricamente ricco, tipologicamente variabile e teoricamente significativo di correlazioni implicazionali robuste e cross-linguisticamente attestate riguarda la posizione relativa dei numerali (parole o morfemi che esprimono quantità numerica discreta come "uno", "due", "tre", "dieci", "cento", "mille", e così via attraverso il sistema di numerazione) rispetto ai nomi che quantificano, modificano e specificano nel sintagma nominale esteso. L'Universale 18 di Greenberg, che si è rivelato essere uno degli universali più robusti, meglio confermati empiricamente, meno problematici e meno soggetti a eccezioni significative nell'inventario originale dei quarantacinque universali proposti nel 1963, afferma con precisione predittiva notevole, chiarezza empirica e semplicità formale che quando i descrittori qualificativi o aggettivi descrittivi precedono il nome manifestando ordine Aggettivo-Nome nella serializzazione lineare superficiale, allora anche i numerali precedono sempre o quasi sempre il nome manifestando ordine Numerale-Nome.

Formalizzato rigorosamente come implicazione logica unidirezionale nella notazione standard della logica proposizionale: Aggettivi_prenominali → Numerali_prenominali, oppure equivalentemente in termini della teoria degli insiemi e delle relazioni di inclusione: L(Agg_prenominali) ⊆ L(Num_prenominali), dove L(Agg_prenominali) denota rigorosamente l'insieme di tutte le lingue naturali umane che possiedono la proprietà di avere aggettivi qualificativi collocati in posizione prenominale come ordine basico, dominante o esclusivo, e L(Num_prenominali) denota corrispondentemente l'insieme di tutte le lingue che possiedono numerali prenominali. L'implicazione formale afferma che l'insieme delle lingue con aggettivi prenominali è un sottoinsieme (proprio o improprio secondo i casi) dell'insieme delle lingue con numerali prenominali, il che equivale logicamente ad affermare che ogni lingua con aggettivi prenominali ha necessariamente anche numerali prenominali, ma non viceversa (esistono lingue con numerali prenominali ma aggettivi postnominali).

In termini di probabilità condizionale quantitativa empiricamente stimabile da campioni tipologici: P(Numerali_prenominali | Aggettivi_prenominali) ≈ 0,98-0,99, rappresentando una delle correlazioni probabilistiche più forti, più robuste e meno soggette a eccezioni sistematiche tra tutte quelle identificate da Greenberg e confermate dalla ricerca tipologica successiva. Questo valore probabilistico estremamente alto, vicinissimo all'unità (certezza assoluta), indica che data una lingua con aggettivi prenominali verificati empiricamente, la probabilità che quella stessa lingua abbia anche numerali prenominali è praticamente universale o quasi-universale, con eccezioni talmente rare da essere statisticamente e tipologicamente trascurabili o marginali.

L'importanza e la profondità teorica di questo universale implicazionale risiede fondamentalmente nel fatto significativo che esso stabilisce una gerarchia implicazionale rigorosa, asimmetrica e cross-linguisticamente robusta tra due tipi distinti e categorialmente differenziabili di modificatori nominali (aggettivi qualificativi versus numerali quantificazionali), suggerendo fortemente e predicendo teoricamente che i numerali occupano una posizione strutturale più "vicina", più nucleare o più strettamente legata sintatticamente e semanticamente alla testa nominale rispetto agli aggettivi descrittivi qualificativi ordinari, i quali sono relativamente più periferici, più opzionali e meno intrinsecamente legati al nucleo referenziale del sintagma nominale.

Questa gerarchia implicazionale può essere interpretata formalmente e teoricamente in termini di una scala di modificazione nominale o gerarchia di proiezione sintattica articolata in multipli livelli strutturali, dove diversi tipi funzionalmente e semanticamente distinti di modificatori, determinanti e specificatori si collocano a diverse altezze strutturali o distanze gerarchiche dalla testa nominale lessicale basica, riflettendo la loro relazione semantica più o meno intrinseca, necessaria e definitoria con le proprietà intensionali, estensionali e referenziali del referente nominale.

Una possibile gerarchia di modificazione nominale, supportata convergentemente da evidenza cross-linguistica indipendente da molteplici domini grammaticali oltre al semplice ordine lineare superficiale (inclusi fenomeni di incorporazione nominale produttiva, patterns diacronici di grammaticalizzazione, ordini tipici di acquisizione nel linguaggio infantile, e restrizioni strutturali sulla coordinazione, ellissi e pronominalizzazione), potrebbe essere plausibilmente strutturata e articolata gerarchicamente come la scala seguente, dove gli

elementi a sinistra sono strutturalmente più vicini o più nucleari rispetto a quelli a destra che sono progressivamente più periferici:

Nome_lessicale < Numerale < Aggettivo_qualificativo < Aggettivo_dimensionale < Aggettivo_valutativo < Dimostrativo < Quantificatore_universale < Articolo_determinativo

dove il simbolo < indica la relazione "più strettamente legato a" o "strutturalmente più vicino a" nella gerarchia di proiezione sintattica estesa. Secondo questa gerarchia articolata e motivata empiricamente, i numerali cardinali sono strutturalmente e semanticamente più vicini al nome lessicale basico dei normali aggettivi descrittivi qualificativi, e quindi sono più conservativi, più resistenti al riposizionamento tipologico e più stabili nel mantenere l'ordine armonico correlato con altri aspetti della struttura nominale e verbale.

Le lingue altaiche del ramo turco, mongolico e tunguso, che sono prototipicamente, armonicamente e sistematicamente SOV con modificatori nominali prevalentemente o esclusivamente prenominali manifestando ordine head-final consistente attraverso multipli domini grammaticali, mostrano consistentemente e senza eccezioni sistematiche numerali categoricamente prenominali in perfetta armonia con gli aggettivi similmente prenominali:

Turco moderno standard: "üç kitap" (tre libro = "tre libri"), "beş öğrenci" (cinque studente = "cinque studenti"), "on ev" (dieci casa = "dieci case"), "yüz insan" (cento persona = "cento persone"). Gli aggettivi qualificativi sono similmente e sistematicamente prenominali: "kırmızı kitap" (rosso libro = "libro rosso"), "büyük ev" (grande casa = "casa grande"), "güzel kız" (bello ragazza = "bella ragazza").

Mongolo khalka (mongolo standard moderno della Mongolia): "gurban nom" (tre libro = "tre libri"), "tawan oyutan" (cinque studente = "cinque studenti"), "arban ger" (dieci casa/tenda = "dieci case"). Aggettivi prenominali: "ulaan nom" (rosso libro = "libro rosso"), "tom ger" (grande casa = "casa grande").

Manciù (lingua tungusa storicamente importante ma ora criticamente endangered): "ilan bithe" (tre libro = "tre libri"), "sunja niyalma" (cinque persona = "cinque persone"). Aggettivi prenominali: "fulgiyan bithe" (rosso libro = "libro rosso").

Le lingue sino-tibetane come il cinese mandarino, il birmano, il tibetano e il newari, che hanno modificatori nominali prevalentemente o sistematicamente prenominali nonostante l'ordine basico controverso e tipologicamente misto SVO del cinese nelle frasi principali, mostrano numerali rigorosamente prenominali:

Cinese mandarino: "sān běn shū" (tre CLASSIFICATORE libro = "tre libri"), "wǔ ge xuéshēng" (cinque CLASSIFICATORE studente = "cinque studenti"), "shí jiān fángzi" (dieci CLASSIFICATORE casa = "dieci case"). Notare l'uso obbligatorio di classificatori nominali (measure words) che si interpongono tra il numerale e il nome, riflettendo la tipologia classificatoria del cinese dove i nomi contabili richiedono classificatori appropriati semanticamente e categorialmente selezionati. Gli aggettivi sono prenominali: "hóng shū" (rosso libro = "libro rosso"), "dà fángzi" (grande casa = "casa grande").

Birmano (birmano standard, lingua ufficiale del Myanmar): "thoun: sa.ouq" (tre libro = "tre libri"), "nga: kyaun:tha:" (cinque studente = "cinque studenti"). Aggettivi prenominali: "a.ni sa.ouq" (rosso libro = "libro rosso"), "kyi: ein" (grande casa = "casa grande").

Tibetano (tibetano standard centrale di Lhasa): "sum teb" (tre libro = "tre libri"), "lnga slob.ma" (cinque studente = "cinque studenti"). Aggettivi prenominali: "dmar.po teb" (rosso libro = "libro rosso"), "che.ba khang.pa" (grande casa = "casa grande").

Le lingue uralico-finniche del ramo finnico e ugrico, che sono sistematicamente SOV o variabilmente SOV/SVO con forte tendenza prenominale per modificatori, mostrano numerali prenominali:

Finlandese (suomi): "kolme kirjaa" (tre libro.PARTITIVO = "tre libri"), "viisi opiskelijaa" (cinque studente.PARTITIVO = "cinque studenti"), "kymmenen taloa" (dieci casa.PARTITIVO = "dieci case"). Notare che i nomi quantificati da numerali maggiori di uno prendono obbligatoriamente il caso partitivo singolare piuttosto che il nominativo o accusativo, riflettendo un pattern morfologico peculiare del finnico. Aggettivi prenominali: "punainen kirja" (rosso libro.NOMINATIVO = "libro rosso"), "suuri talo" (grande casa = "casa grande").

Estone (eesti keel): "kolm raamatut" (tre libro.PARTITIVO = "tre libri"), "viis üliõpilast" (cinque studente.PARTITIVO = "cinque studenti"). Aggettivi prenominali: "punane raamat" (rosso libro = "libro rosso"), "suur maja" (grande casa = "casa grande").

Ungherese (magyar): "három könyv" (tre libro.NOMINATIVO = "tre libri"), "öt diák" (cinque studente = "cinque studenti"), "tíz ház" (dieci casa = "dieci case"). Aggettivi prenominali: "piros könyv" (rosso libro = "libro rosso"), "nagy ház" (grande casa = "casa grande").

Questa consistenza impressionante, robusta e sistematica del pattern numerali prenominali attraverso lingue geneticamente non correlate, filogeneticamente distanti e geograficamente distribuite su continenti diversi (lingue altaiche in Asia centrale e settentrionale, lingue sino-tibetane in Asia orientale e sudorientale, lingue uralico-finniche in Europa settentrionale e centrale, più numerose altre famiglie non qui elencate per brevità), tutte convergenti tipologicamente verso la stessa configurazione strutturale, fornisce evidenza empirica molto forte, statisticamente robusta e metodologicamente convincente per la validità dell'Universale 18 come generalizzazione tipologica genuina, cross-linguisticamente significativa e teoricamente profonda piuttosto che come artefatto superficiale o correlazione accidentale limitata a una singola area geografica o famiglia genetica.

Il pattern complementare, inverso e speculare si osserva nelle lingue con aggettivi postnominali o con ordini misti, dove i numerali possono manifestare sia posizione prenominale sia posizione postnominale, mostrando maggiore variabilità tipologica, flessibilità strutturale e dipendenza da fattori linguistici addizionali. Tuttavia, criticamente e crucialmente per la validità dell'universale implicazionale, questo pattern asimmetrico (Aggettivi_prenominali implica fortemente e quasi-universalmente Numerali_prenominali, ma Aggettivi_postnominali non implica altrettanto fortemente e deterministicamente Numerali_postnominali, permettendo entrambe le configurazioni) riflette precisamente e conferma la posizione gerarchica proposta dei numerali come modificatori strutturalmente più nucleari, semanticamente più intrinseci e tipologicamente più conservativi rispetto agli aggettivi qualificativi ordinari.

Le lingue romanze manifestano pattern interessanti, variabili e teoricamente illuminanti che confermano l'asimmetria fondamentale dell'universale. L'italiano standard contemporaneo ha generalmente e prevalentemente aggettivi postnominali come ordine basico non marcato per la maggior parte delle classi semantiche di aggettivi: "libro rosso" (libro rosso), "casa grande" (casa grande), "ragazzo intelligente" (ragazzo intelligente), "città bella" (città bella), con l'aggettivo che segue il nome testa. Tuttavia, alcuni aggettivi valutativi, dimensionali o altamente frequenti possono o devono precedere il nome con differenze semantiche sottili o pragmatiche: "bel libro" versus "libro bello", "gran casa" versus "casa grande", "buon vino" versus "vino buono", dove l'ordine prenominale tende ad avere interpretazione più valutativa, soggettiva o idiomatica mentre l'ordine postnominale è più descrittivo, oggettivo e compositivo.

I numerali cardinali, criticamente e crucialmente per la validità dell'Universale 18, sono categoricamente, rigidamente e senza alcuna eccezione o variabilità prenominali in italiano: "tre libri", "cinque studenti", "dieci case", "cento persone", "mille euro", senza alcuna possibilità grammaticale di inversione (*"libri tre" o *"studenti cinque" sono completamente agrammaticali e impossibili in italiano standard, sebbene possano occasionalmente apparire in poesia, titoli stilisticamente marcati o trascrizioni di varietà dialettali con influenze substratali diverse).

Questo pattern conferma perfettamente e senza ambiguità l'Universale 18: anche quando l'italiano permette variazione significativa e semanticamente condizionata nella posizione degli aggettivi qualificativi con alcuni aggettivi prenominali e altri postnominali, i numerali cardinali rimangono invariabilmente e rigidamente fissi in posizione prenominale, riflettendo la loro posizione strutturale più nucleare e la loro resistenza maggiore al riposizionamento tipologico.

Il francese mostra pattern molto simile e strutturalmente parallelo: aggettivi prevalentemente postnominali come ordine non marcato ("livre rouge", "maison grande", "garçon intelligent"), ma alcuni aggettivi valutativi, dimensionali o altamente frequenti prenominali ("beau livre", "grande maison", "bon vin", "petit chat", "jeune fille"), mentre i numerali sono categoricamente e invariabilmente prenominali: "trois livres", "cinq étudiants", "dix maisons", "cent personnes", "mille euros", mai *"livres trois".

Lo spagnolo: aggettivi generalmente postnominali ("libro rojo", "casa grande", "chico inteligente"), alcuni prenominali con valore valuativo o intensificato ("buen libro", "gran casa"), numerali categoricamente prenominali: "tres libros", "cinco estudiantes", "diez casas", "cien personas", "mil euros".

Il portoghese: pattern identico con aggettivi prevalentemente postnominali ("livro vermelho", "casa grande"), alcuni prenominali ("bom livro", "grande casa"), numerali prenominali: "três livros", "cinco estudantes", "dez casas", "cem pessoas", "mil euros".

Il rumeno: aggettivi tipicamente postnominali con alcuni prenominali, numerali prenominali: "trei cărți" (tre libri), "cinci studenți" (cinque studenti).

L'inglese, con aggettivi rigidamente e categoricamente prenominali senza eccezioni significative ("red book", "big house", "intelligent boy", mai *"book red" o *"house big" che sono completamente agrammaticali), ha naturalmente e predicibilmente numerali similmente

prenominali in perfetta armonia strutturale: "three books", "five students", "ten houses", "hundred people", "thousand euros", conformandosi perfettamente all'Universale 18 nella sua forma implicazionale più diretta.

Il tedesco mostra pattern identico e strutturalmente parallelo: aggettivi rigidamente prenominali con morfologia flessiva complessa di accordo ("rotes Buch" nominativo neutro, "rotem Buch" dativo neutro, "roter Bücher" genitivo plurale, "großes Haus", "intelligenter Junge"), numerali prenominali: "drei Bücher", "fünf Studenten", "zehn Häuser", "hundert Personen", "tausend Euro".

Le lingue celtiche come il gallese (Cymraeg), l'irlandese (Gaeilge), il bretone (Brezhoneg) e lo scozzese gaelico (Gàidhlig), che sono tipologicamente VSO con ordine head-initial e manifestano conseguentemente aggettivi tipicamente postnominali in armonia con questo ordine generale, mostrano interessantemente numerali prenominali, costituendo un'apparente eccezione parziale all'armonia direzionale stretta ma confermando criticamente e perfettamente l'Universale 18 nella sua forma implicazionale precisa.

Gallese: aggettivi rigorosamente postnominali con mutazione consonantica iniziale (soft mutation) del nome testa ma non dell'aggettivo: "llyfr coch" (libro rosso, con llyfr < lyfr dopo articolo), "tŷ mawr" (casa grande), "bachgen clyfar" (ragazzo intelligente). Numerali prenominali con mutazioni consonantiche complesse governate dal numerale: "tri llyfr" (tre libro, con soft mutation), "pump myfyriwr" (cinque studente, con nasal mutation), "deg tŷ" (dieci casa).

Irlandese: aggettivi postnominali: "leabhar dearg" (libro rosso), "teach mór" (casa grande), "buachaill cliste" (ragazzo intelligente). Numerali prenominali con mutazioni consonantiche e forme speciali: "trí leabhar" (tre libro), "cúig mhac léinn" (cinque studente, con lenition), "deich dteach" (dieci casa, con eclipsis).

Questo pattern nelle lingue celtiche è teoricamente molto significativo e empiricamente cruciale perché dimostra convincentemente che i numerali occupano una posizione strutturale categorialmente distinta, gerarchicamente separata e funzionalmente indipendente dagli aggettivi descrittivi ordinari nel sintagma nominale esteso, e sono soggetti a principi di collocazione posizionale parzialmente autonomi e indipendenti dagli aggettivi. Alcune analisi sintattiche formali hanno proposto che i numerali in queste lingue, e più generalmente cross-linguisticamente, non sono genuini modificatori aggettivali attributivi ma piuttosto determinanti quantificazionali che occupano una proiezione funzionale distinta, strutturalmente più alta e gerarchicamente più esterna nella struttura articolata del sintagma nominale (probabilmente una proiezione NumP o QP situata tra DP e NP nella cartografia fine del sintagma nominale), più vicina strutturalmente ai determinanti definitivi e ai dimostrativi che agli aggettivi descrittivi qualificativi basici.

Questa analisi strutturale alternativa predicherebbe naturalmente e spiegherebbe teoricamente che i numerali si comportino sintatticamente e posizionalmente in modo diverso, indipendente e parzialmente autonomo dagli aggettivi qualificativi per quanto riguarda la posizione lineare, le mutazioni fonologiche trigger, le restrizioni di co-occorrenza e altri fenomeni grammaticali, proprio perché occupano posizioni strutturali diverse nella gerarchia sintattica estesa del sintagma nominale e appartengono a categorie funzionali distinte piuttosto che alla stessa categoria lessicale.

Una conseguenza interessante e teoricamente profonda dell'Universale 18 riguarda le gerarchie di accessibilità tipologica, la marcatezza relativa di diverse configurazioni e le implicazioni per l'acquisizione linguistica, il cambiamento diacronico e la processabilità cognitiva. Se definiamo quattro configurazioni logicamente possibili basate sulla combinazione binaria ortogonale di posizione di aggettivi e numerali: (1) Aggettivi_prenominali combinati con Numerali_prenominali, (2) Aggettivi_prenominali combinati con Numerali_postnominali, (3) Aggettivi_postnominali combinati con Numerali_prenominali, (4) Aggettivi_postnominali combinati con Numerali_postnominali, l'Universale 18 prevede esplicitamente e formalmente che la configurazione 2 (aggettivi prenominali con numerali postnominali) sia estremamente rara, marcatissima o virtualmente inesistente nella tipologia mondiale documentata.

Evidenza empirica sistematica da database tipologici ampi e metodologicamente controllati conferma questa previsione teorica con notevole robustezza statistica e consistenza cross-linguistica. Su un campione rappresentativo di circa 1200 lingue con informazione completa e affidabile su entrambe le variabili critiche nel World Atlas of Language Structures (WALS Online, edizione 2013 e aggiornamenti successivi), la distribuzione approssimativa osservata è:

- Configurazione 1 (Aggettivi_prenominali + Numerali_prenominali): circa 450 lingue, rappresentanti il 37,5% del campione totale
- Configurazione 3 (Aggettivi_postnominali + Numerali_prenominali): circa 400 lingue, rappresentanti il 33,3% del campione
- Configurazione 4 (Aggettivi_postnominali + Numerali_postnominali): circa 340 lingue, rappresentanti il 28,3% del campione
- Configurazione 2 (Aggettivi_prenominali + Numerali_postnominali): circa 10 lingue, rappresentanti solo lo 0,8% del campione

La configurazione 2, quella prevista come impossibile o estremamente marcata dall'Universale 18, rappresenta infatti meno dell'1% del campione totale, confermando impressionantemente che è una configurazione estremamente marcata, altamente eccezionale e tipologicamente marginale. Le pochissime lingue che sembrano mostrare apparentemente questo pattern disarmonico e violativo dell'universale sono tipicamente casi metodologicamente problematici, empiricamente ambigui o analiticamente controversi: lingue con ordini multipli variabili e condizionati da fattori pragmatici, semantici o stilistici complessi; lingue in transizione tipologica attiva con ordini instabili e generazionalmente variabili; o lingue con influenze areali complesse da contatto linguistico prolungato e intenso con lingue tipologicamente diverse, portando a stratificazione di pattern eterogenei.

Possiamo calcolare quantitativamente la forza dell'implicazione mediante la probabilità condizionale: P(Numerali_prenominali | Aggettivi_prenominali) = 450 / (450 + 10) = 450 / 460 ≈ 0,978, una probabilità condizionale estremamente alta vicinissima all'unità che conferma la robustezza empirica quasi-universale dell'universale implicazionale. Al contrario, l'implicazione inversa ipotetica è molto più debole e sostanzialmente non significativa: P(Aggettivi_prenominali | Numerali_prenominali) = 450 / (450 + 400) = 450 / 850 ≈ 0,529, solo marginalmente e non significativamente sopra 0,50 (assenza di correlazione), confermando quantitativamente l'asimmetria fondamentale e irriducibile della relazione implicazionale: gli aggettivi prenominali implicano fortemente numerali prenominali, ma i

numerali prenominali non implicano necessariamente aggettivi prenominali, proprio come predetto teoricamente dalla gerarchia di nuclearità strutturale proposta.

Comparativi e Superlativi: L'Universale 22 e la Gerarchia Implicazionale nella Marcatura del Grado

Un dominio grammaticale addizionale e teoricamente significativo dove Greenberg identifica correlazioni implicazionali robuste, cross-linguisticamente attestate e tipologicamente rilevanti riguarda la formazione morfologica e sintattica dei gradi di comparazione degli aggettivi qualificativi, ovvero le costruzioni grammaticali specializzate che esprimono il grado relativo, comparativo o superlativo di una proprietà o qualità descritta dall'aggettivo. Le lingue del mondo mostrano fondamentalmente due strategie principali, tipologicamente distinte e morfosintatticamente contrastive, per formare e codificare i gradi comparativi e superlativi degli aggettivi:

Prima strategia, morfologica o sintetica: il comparativo e il superlativo sono formati mediante affissazione derivazionale diretta, modificazione morfologica interna (come umlaut, ablaut o altri processi morfofonologici) o supplettivismo lessicale del tema aggettivale basico, producendo forme flesse o derivate che codificano morfologicamente e unitariamente il significato comparativo o superlativo. Esempi rappresentativi da diverse lingue includono: inglese "tall" (alto) → "taller" (più alto, comparativo mediante suffissazione -er) → "tallest" (altissimo/il più alto, superlativo mediante suffissazione -est); latino classico "altus" (alto) → "altior" (più alto, comparativo mediante suffisso -ior/-ius) → "altissimus" (altissimo, superlativo mediante suffisso -issimus); tedesco standard "groß" (grande) → "größer" (più grande, comparativo con umlaut vocalico e suffisso -er) → "größt-" (il più grande, superlativo con umlaut e suffisso -st più flessione); russo "krasivyj" (красивый, bello) → "krasiveje" (красивее, più bello, comparativo sintetico) → "krasivejšij" (красивейший, bellissimo, superlativo sintetico).

Seconda strategia, analitica o perifrastica: il comparativo e il superlativo sono formati mediante parole grammaticali separate e indipendenti fonologicamente, tipicamente avverbi intensificatori, particelle comparativizzanti o marcatori funzionali di grado che si combinano sintatticamente con la forma basica non derivata dell'aggettivo, producendo costruzioni sintattiche multi-parola che codificano analiticamente il significato comparativo o superlativo. Esempi rappresentativi includono: inglese "beautiful" (bello) → "more beautiful" (più bello, comparativo analitico con avverbio "more") → "most beautiful" (bellissimo/il più bello, superlativo analitico con avverbio "most"); francese moderno "beau" (bello) → "plus beau" (più bello, comparativo con avverbio "plus") → "le plus beau" (il più bello, superlativo con articolo determinativo e avverbio intensificatore "plus"); italiano standard "bello" → "più bello" (comparativo analitico) → "il più bello" (superlativo relativo analitico) o "bellissimo" (superlativo assoluto sintetico mediante suffisso -issimo); cinese mandarino "hǎo" (好, buono) → "gèng hǎo" (更好, più buono, comparativo con avverbio "gèng") → "zuì hǎo" (最好, il più buono, superlativo con avverbio "zuì"); vietnamita "đẹp" (bello) → "đẹp hơn" (bello più = più bello) → "đẹp nhất" (bello più.superlativo = bellissimo).

L'Universale 22 di Greenberg, formulato con precisione e semplicità caratteristiche, afferma che se una lingua forma il grado comparativo mediante suffissazione morfologica diretta

dell'aggettivo (strategia morfologica sintetica), allora quella stessa lingua forma necessariamente o quasi sempre anche il grado superlativo mediante suffissazione morfologica; in altre parole, se una lingua adotta una strategia morfologica sintetica per codificare il comparativo, adotterà coerentemente e sistematicamente la stessa strategia morfologica sintetica anche per codificare il superlativo, manifestando consistenza tipologica nella scelta della strategia di codifica del grado attraverso i due livelli della scala comparativa.

Formalmente espresso come implicazione logica unidirezionale: Comparativo_morfologico → Superlativo_morfologico, oppure in notazione insiemistica: L(Comp_morf) ⊆ L(Sup_morf), dove L(Comp_morf) denota l'insieme delle lingue che formano il comparativo morfologicamente mediante affissazione e L(Sup_morf) denota l'insieme delle lingue che formano il superlativo morfologicamente. In termini probabilistici quantitativi: P(Superlativo_morfologico | Comparativo_morfologico) ≈ 0,90-0,95, indicando che data una lingua con comparativo morfologico, la probabilità che abbia anche superlativo morfologico è molto alta, dell'ordine del 90-95%.

Questo universale implicazionale stabilisce una gerarchia nella marcatura del grado che può essere formalizzata come: Superlativo > Comparativo, dove il simbolo > indica "più marcato di" o "più soggetto a strategia morfologica esplicita di". L'interpretazione teorica più plausibile e comunemente adottata è che il superlativo, essendo semanticamente il grado più estremo, più complesso intensionalmente e più ristretto estensionalmente (richiede confronto esaustivo con l'intero insieme di riferimento rilevante piuttosto che semplice confronto binario con un singolo comparando specifico come il comparativo), sia tipologicamente più marcato, acquisizionalmente più difficile, diacronicamente più instabile e quindi più suscettibile di ricevere marcatura morfologica esplicita, specializzata e dedicata. Se una lingua investe risorse morfologiche per marcare esplicitamente il grado relativamente meno marcato (comparativo mediante affissazione), allora a fortiori e necessariamente investirà risorse morfologiche anche per il grado più marcato (superlativo), ma non necessariamente viceversa.

L'inglese contemporaneo standard fornisce un esempio paradigmatico e metodologicamente illuminante, anche se parzialmente complicato dalla coesistenza sincronicamente attestata di strategie morfologiche e analitiche condizionate lessicalmente e fonologicamente dalla lunghezza sillabica e dalla struttura morfologica dell'aggettivo basico. Per aggettivi monosillabici e molti aggettivi bisillabici (particolarmente quelli terminanti in -y, -le, -ow, -er), l'inglese usa sistematicamente e produttivamente suffissazione morfologica per entrambi i gradi: "tall" (alto) → "taller" (comparativo mediante suffisso -er) → "tallest" (superlativo mediante suffisso -est); "big" (grande) → "bigger" (con geminazione consonantica) → "biggest"; "happy" (felice) → "happier" (con cambio ortografico y→i) → "happiest"; "simple" (semplice) → "simpler" → "simplest"; "narrow" (stretto) → "narrower" → "narrowest"; "clever" (intelligente) → "cleverer" → "cleverest".

Per aggettivi polisillabici più lunghi (tipicamente trisillabici o maggiori, e molti bisillabici che non rientrano nelle classi morfologiche privilegiate), l'inglese usa strategia analitica perifrastica con gli avverbi intensificatori "more" e "most": "beautiful" (bello) → "more beautiful" (comparativo analitico) → "most beautiful" (superlativo analitico); "interesting" (interessante) → "more interesting" → "most interesting"; "intelligent" (intelligente) → "more

intelligent" → "most intelligent"; "comfortable" (comodo) → "more comfortable" → "most comfortable".

Criticamente e crucialmente per la validità dell'Universale 22, l'inglese non presenta aggettivi che formano il comparativo morfologicamente mediante suffissazione ma il superlativo analiticamente mediante perifrasi (*"taller" ma "most tall" è fortemente agrammaticale, marcato o al massimo marginalmente accettabile in contesti estremamente specifici e marcati, non rappresentando il pattern produttivo normale). Al contrario, la combinazione inversa ipotetica dove il comparativo è formato analiticamente ma il superlativo morfologicamente è altrettanto impossibile o estremamente marginale (*"more tall" con "tallest" suona anomalo e disarmonico quando "more tall" non è la forma standard accettata nella varietà normativa prescrittiva).

Questo pattern di coerenza strategica conferma l'Universale 22: per ogni singolo aggettivo lessicale individuale, la strategia morfologica versus analitica è consistente attraverso i due gradi comparativo e superlativo, senza mescolanze disarmoniche dove un grado usa strategia morfologica e l'altro strategia analitica per lo stesso aggettivo. La scelta della strategia (morfologica versus analitica) è determinata lessicalmente e fonologicamente dall'aggettivo basico (lunghezza sillabica, struttura morfologica, frequenza d'uso), ma una volta determinata questa scelta, essa si applica coerentemente a entrambi i gradi.

Il latino classico, lingua indoeuropea arcaica e altamente conservativa con morfologia nominale e aggettivale estremamente ricca, articolata e sintetica, mostra il pattern con particolare chiarezza cristallina, regolarità morfologica e trasparenza paradigmatica, essendo una lingua quasi esclusivamente e categoricamente morfologica nella formazione dei gradi comparativi e superlativi senza ricorso significativo a strategie analitiche perifrastiche. Il latino forma comparativi regolari mediante il suffisso flessivo -ior (maschile/femminile) / -ius (neutro) aggiunto al tema aggettivale: "altus, -a, -um" (alto) → "altior, altius" (più alto, comparativo con flessione per genere); "fortis, -e" (forte) → "fortior, fortius" (più forte); "sapiens, sapientis" (saggio) → "sapientior, sapientius" (più saggio); "felix, felicis" (felice) → "felicior, felicius" (più felice).

I superlativi regolari sono formati mediante il suffisso flessivo -issimus, -issima, -issimum (con accordo completo per genere, numero e caso secondo le declinazioni nominali): "altus" → "altissimus, -a, -um" (altissimo, il più alto); "fortis" → "fortissimus, -a, -um" (fortissimo); "sapiens" → "sapientissimus" (sapientissimo); "felix" → "felicissimus" (felicissimo). La strategia morfologica sintetica è applicata coerentemente, sistematicamente e produttivamente a entrambi i gradi senza eccezioni significative per aggettivi regolari, con variazioni morfologiche limitate a pochi aggettivi suppletivi altamente frequenti ("bonus, melior, optimus" = buono, migliore, ottimo; "malus, peior, pessimus" = cattivo, peggiore, pessimo; "magnus, maior, maximus" = grande, maggiore, massimo; "parvus, minor, minimus" = piccolo, minore, minimo) che comunque mantengono la strategia morfologica sintetica anche se con radici suppletive diverse.

Il tedesco moderno standard (Hochdeutsch) manifesta strategia morfologica sintetica coerente e altamente produttiva per la stragrande maggioranza degli aggettivi qualificativi comuni e frequenti, con pattern regolari di suffissazione combinati frequentemente con modificazioni morfofonologiche interne (umlaut vocalico) per aggettivi monosillabici: "groß"

(grande) → "größer" (comparativo con umlaut ö ← o e suffisso -er) → "größt-" (superlativo con umlaut e suffisso -st- più flessioni casuali e di genere secondo l'uso attributivo o predicativo); "klein" (piccolo) → "kleiner" → "kleinst-"; "alt" (vecchio) → "älter" (con umlaut ä ← a) → "ältest-"; "jung" (giovane) → "jünger" (con umlaut ü ← u) → "jüngst-"; "schön" (bello) → "schöner" → "schönst-"; "dumm" (stupido) → "dümmer" → "dümmst-".

Anche aggettivi relativamente lunghi, polisillabici e morfologicamente complessi usano preferenzialmente suffissazione sintetica piuttosto che perifrasi analitica, a differenza dell'inglese che restringe la strategia morfologica principalmente agli aggettivi corti: "interessant" (interessante) → "interessanter" → "interessantest-"; "intelligent" (intelligente) → "intelligenter" → "intelligentest-"; "kompliziert" (complicato) → "komplizierter" → "komplizierteste-". La strategia morfologica sintetica è applicata coerentemente attraverso entrambi i gradi per ogni aggettivo, confermando perfettamente l'Universale 22.

Le lingue romanze moderne (italiano, francese, spagnolo, portoghese, rumeno, catalano, occitano), al contrario del loro antenato latino classico morfologicamente ricco, hanno largamente abbandonato progressivamente attraverso l'evoluzione diacronica la strategia morfologica sintetica ereditata dal latino per i comparativi (perdendo completamente i suffissi comparativi latini -ior/-ius) e hanno adottato strategie prevalentemente o esclusivamente analitiche perifrastiche per la formazione dei comparativi usando avverbi intensificatori grammaticalizzati derivanti da aggettivi latini di quantità ("più" ← lat. PLUS, "más" ← MAGIS, "plus" ← PLUS, "mais" ← MAGIS).

Tuttavia, le lingue romanze mantengono parzialmente la strategia morfologica sintetica per i superlativi assoluti (che esprimono grado molto alto della qualità senza confronto esplicito con un insieme di riferimento) attraverso il suffisso derivazionale produttivo -issimo/-íssimo/-issime ereditato direttamente e conservativamente dal latino -issimus, mentre usano strategie analitiche per i superlativi relativi (che esprimono il grado massimo in confronto con un insieme definito di entità).

L'italiano standard contemporaneo forma comparativi esclusivamente e categoricamente con la strategia analitica usando l'avverbio intensificatore "più" + aggettivo basico non modificato, e forma superlativi usando due strategie parallele e coesistenti: superlativo relativo analitico con articolo determinativo + "più" + aggettivo ("il/la/i/le più" + aggettivo), e superlativo assoluto sintetico con suffisso -issimo/-issima/-issimi/-issime: "bello, bella" (bello) → "più bello/bella" (comparativo analitico, più bello) → "il più bello / la più bella" (superlativo relativo analitico, il più bello) oppure "bellissimo/bellissima" (superlativo assoluto sintetico, bellissimo).

Altri esempi italiani: "grande" → "più grande" → "il più grande" / "grandissimo"; "intelligente" → "più intelligente" → "il più intelligente" / "intelligentissimo"; "veloce" → "più veloce" → "il più veloce" / "velocissimo"; "importante" → "più importante" → "il più importante" / "importantissimo". La strategia è coerente: comparativo sempre analitico, superlativo relativo sempre analitico armonicamente, superlativo assoluto sintetico come strategia addizionale disponibile e altamente produttiva.

Anche le lingue romanze confermano l'Universale 22 perché, sebbene abbiano abbandonato la strategia morfologica per i comparativi a favore della strategia analitica (trasformazione diacronica significativa rispetto al latino), esse adottano strategie coerenti per entrambi i tipi

di superlativo: quando usano strategia analitica per il comparativo ("più bello"), usano strategia analitica coerentemente anche per il superlativo relativo ("il più bello"), manifestando consistenza tipologica; e quando usano strategia morfologica per il superlativo assoluto ("-issimo"), questa è una strategia morfologica addizionale, parallela e coesistente oltre a quella analitica, non in sostituzione esclusiva.

Criticamente per la validità dell'universale, non esistono lingue romanze che formano il comparativo morfologicamente mediante suffissazione ma il superlativo analiticamente mediante perifrasi, né viceversa per lo stesso aggettivo. La transizione diacronica dal latino alle lingue romanze ha coinvolto un cambiamento sistemico e coerente da strategia prevalentemente morfologica (latino: altior, altissimus, entrambi sintetici) a strategia prevalentemente analitica per comparativi e superlativi relativi (italiano: più alto, il più alto, entrambi analitici), mantenendo tuttavia coerenza strategica attraverso il cambiamento.

Il francese moderno standard: "beau" → "plus beau" (comparativo analitico) → "le plus beau" (superlativo relativo analitico) / "très beau" (intensificazione analitica, non propriamente superlativo assoluto morfologico produttivo come l'italiano -issimo). Lo spagnolo: "hermoso" → "más hermoso" → "el más hermoso" / "hermosísimo" (superlativo assoluto sintetico altamente produttivo). Il portoghese: "bonito" → "mais bonito" → "o mais bonito" / "bonitíssimo". Il rumeno: "frumos" → "mai frumos" → "cel mai frumos".

Le lingue altaiche mostrano strategie morfologiche o semi-analitiche coerenti con marcatori dedicati di grado che si comportano morfosintatticamente in modo intermedio tra affissi flessivi puri e particelle grammaticali indipendenti. Il turco moderno forma comparativi usando l'avverbio grammaticalizzato "daha" (più, originariamente da persiano) che precede l'aggettivo basico, e forma superlativi usando l'avverbio grammaticalizzato "en" (il più) che similmente precede l'aggettivo: "büyük" (grande) → "daha büyük" (più grande, comparativo) → "en büyük" (il più grande, superlativo); "güzel" (bello) → "daha güzel" → "en güzel"; "iyi" (buono) → "daha iyi" → "en iyi"; "küçük" (piccolo) → "daha küçük" → "en küçük".

Sebbene queste siano tecnicamente strategie semi-analitiche o perifrastiche piuttosto che puramente morfologiche mediante affissazione flessiva diretta, esse mostrano coerenza sistemica nella strategia di formazione: entrambi i gradi (comparativo e superlativo) usano particelle grammaticalizzate specializzate che precedono l'aggettivo secondo il pattern generale head-final del turco, piuttosto che perifrasi lessicali complete con verbi o costruzioni sintattiche complesse. La coerenza strategica è mantenuta perfettamente.

Il mongolo khalka mostra pattern simili con particelle dedicati: "том" (tom, grande) → "илүү том" (ilüü tom, più grande, con particella comparativa "ilüü") → "хамгийн том" (khamgiin tom, il più grande, con particella superlativa "khamgiin"). Il giapponese usa strategie analitiche coerenti: "大きい" (ōkii, grande) → "もっと大きい" (motto ōkii, più grande, con avverbio "motto") → "一番大きい" (ichiban ōkii, il più grande, con avverbio superlativo "ichiban" letteralmente "numero uno").

Un aspetto teoricamente interessante, tipologicamente significativo e cross-linguisticamente osservabile riguarda la relazione sistematica tra la strategia di formazione preferita dei comparativi/superlativi (morfologica sintetica versus analitica perifrastica) e altre proprietà tipologiche generali della lingua, in particolare il parametro fondamentale di sinteticità versus analiticità della morfologia nominale e verbale complessiva. Le lingue con morfologia

flessiva ricca, complessa e altamente sintetica (misurata quantitativamente mediante l'indice di sintesi, definito come il numero medio di morfemi per parola forme in testi naturali rappresentativi) tendono sistematicamente, robustatamente e predicibilmente a preferire strategie morfologiche sintetiche per i gradi di comparazione, mentre le lingue con morfologia flessiva impoverita, semplificata e altamente analitica (con indice di sintesi basso vicino a 1,0) tendono a preferire strategie analitiche perifrastiche.

Questa correlazione tipologica può essere formalizzata quantitativamente come probabilità condizionale: P(Comparativo_morfologico | Morfologia_sintetica) > P(Comparativo_morfologico | Morfologia_analitica), dove Morfologia_sintetica indica lingue con alto indice di sintesi (numero medio di morfemi per parola > 2,5, caratteristico di lingue agglutinanti come turco, finlandese, swahili, o lingue flessive ricche come latino, russo, sanscrito) e Morfologia_analitica indica lingue con basso indice di sintesi (numero medio di morfemi per parola < 1,5, caratteristico di lingue isolanti come cinese mandarino, vietnamita, thai, o lingue con morfologia ridotta come inglese, francese moderno, afrikaans).

Stime quantitative basate su campioni tipologici ampi e geneticamente diversificati suggeriscono valori approssimativi: P(Comp_morf | Morf_sint) ≈ 0,75 versus P(Comp_morf | Morf_anal) ≈ 0,25, una differenza sostanziale e statisticamente significativa di 50 punti percentuali (0,75 - 0,25 = 0,50) che riflette tendenze generali forti di coerenza tipologica cross-categoriale: le lingue tendono ad essere consistenti nella scelta tra strategie morfologiche versus analitiche attraverso diversi domini grammaticali (flessione nominale casuale, flessione verbale tempo-aspettuale, formazione comparativi/superlativi, formazione diminutivi/accrescitivi, derivazione nominale e verbale, etc.), sebbene eccezioni e pattern misti siano certamente attestati e documentati.

Questa tendenza alla coerenza tipologica trans-dominio può essere teoricamente spiegata attraverso diversi meccanismi causali potenzialmente sinergici e complementari: primo, vincoli cognitivi generali sulla complessità computazionale del processamento morfologico, dove mantenere un tipo uniforme di strategia morfologica riduce il carico cognitivo computazionale; secondo, dinamiche di apprendibilità e acquisizione, dove consistenza tipologica facilita l'acquisizione del sistema grammaticale da parte dei bambini attraverso regolarità e predicibilità aumentate; terzo, processi di cambiamento linguistico diacronico sistemici, dove innovazioni morfologiche o semplificazioni tendono a diffondersi analogicamente attraverso domini grammaticali correlati; quarto, pressioni funzionali comunicative verso economicità espressiva coerente.

I Casi Grammaticali e l'Ordine delle Parole: Universale 41 e la Marcatura Differenziale degli Argomenti

Un'area particolarmente ricca, complessa e teoricamente centrale di universali implicazionali robustamente attestati e profondamente significativi riguarda l'espressione morfologica e sintattica dei casi grammaticali (marcatura flessiva sui sintagmi nominali indicante funzioni sintattiche e ruoli semantici) e dei ruoli tematici o semantici associati agli argomenti nominali delle predicazioni verbali. I casi grammaticali sono marcature morfologiche sistematiche sui nomi, pronomi, e loro modificatori e determinanti concordanti (aggettivi, numerali, articoli, dimostrativi) che indicano esplicitamente e morfologicamente la funzione sintattica che il

sintagma nominale svolge nella struttura della frase (soggetto, oggetto diretto, oggetto indiretto, complemento obliquo, etc.) oppure il ruolo semantico o tematico associato (agente, paziente, esperiente, beneficiario, strumento, locazione, origine, meta, etc.).

Alcune lingue, come il latino classico (con 6 casi morfologicamente distinti: nominativo, genitivo, dativo, accusativo, ablativo, vocativo), il greco antico (5 casi: nominativo, genitivo, dativo, accusativo, vocativo), il russo moderno (6 casi), il finlandese contemporaneo (15 casi morfologicamente produttivi e distintivi), il turco moderno (6 casi basici più casi locativi estesi), il giapponese (sistema ricco di particelle casuali postposizionali), l'ungherese (circa 18-20 casi secondo analisi), e molte altre lingue con morfologia nominale ricca, complessa e altamente sintetica, hanno sistemi di casi estremamente elaborati, articolati e finemente differenziati con molte distinzioni morfologiche esplicite che codificano fino a quindici, venti o anche più relazioni grammaticali e semantiche diverse morfologicamente attraverso affissi flessivi dedicati.

Al contrario, altre lingue come l'inglese moderno (che ha perso quasi completamente la flessione casuale nominale, mantenendo distinzioni casuali visibili morfologicamente solo in alcuni pronomi personali: I/me, he/him, she/her, we/us, they/them, who/whom), il cinese mandarino (lingua isolante priva completamente di morfologia flessiva nominale), il francese moderno (sistema casuale nominale completamente atrofizzato anche nei pronomi in molte varietà parlate colloquiali), lo spagnolo e l'italiano contemporanei (casi nominali completamente persi eccetto tracce pronominali residuali), e molte lingue isolanti, analitiche o con morfologia flessiva impoverita, hanno perso quasi completamente o completamente la flessione nominale casuale originaria ereditata da stadi più antichi e ricorrono invece primariamente e sistematicamente all'ordine fisso rigido delle parole e alle adposizioni (preposizioni o postposizioni) analitiche indipendenti per indicare disambiguare e codificare le relazioni grammaticali e semantiche tra argomenti verbali e predicati.

L'Universale 41 di Greenberg stabilisce una correlazione importante, robusta e teoricamente significativa tra il tipo e l'estensione del sistema di marcatura casuale morfologica e l'ordine basico dei costituenti maggiori verbali. Nella sua formulazione originale precisa, esso afferma che se in una lingua la distinzione fondamentale nominativo-accusativo (la distinzione morfologica basica e primaria tra il caso del soggetto intransitivo/transitivo e il caso dell'oggetto diretto transitivo in lingue con allineamento sintattico nominativo-accusativo) è marcata morfologicamente, codificata esplicitamente e differenziata visibilmente solo nel sistema pronominale (pronomi personali, dimostrativi, interrogativi, relativi) e non estensivamente nel sistema dei nomi comuni lessicali pieni, allora la lingua è quasi certamente e predicibilmente di tipo VO (verbo precede oggetto, includendo tanto SVO quanto VSO come sottotipi).

Questo universale identifica e codifica un'asimmetria tipologica interessante, robusta e teoricamente profonda: le lingue che hanno una marcatura casuale morfologica ridotta, impoverita o selettivamente atrofizzata, limitata morfologicamente ed estensionalmente ai pronomi personali e dimostrativi senza estendersi produttivamente e sistematicamente ai nomi comuni lessicali pieni, tendono sistematicamente, robustatamente e predicibilmente ad essere lingue VO con ordine relativamente fisso e rigido dei costituenti maggiori, suggerendo plausibilmente che l'ordine VO fisso (verbo prima dell'oggetto con conseguente ordine relativamente rigido e predicibile SVO o VSO senza scrambling libero o permutazioni

pragmaticamente condizionate) permette, facilita, rende possibile o forse addirittura favorisce evolutivamente una riduzione progressiva e semplificazione diacronica della marcatura morfologica esplicita del caso nominale, perché la posizione strutturale fissa, rigida e predicibile del costituente nominale nella serializzazione lineare superficiale può indicare sufficientemente, univocamente e non ambiguamente la sua funzione sintattica attraverso criteri puramente posizionali e configurazionali senza necessità assoluta di marcatura morfologica ridondante e sovrabbondante.

Al contrario e complementarmente, le lingue OV, specialmente quelle che permettono ordine dei costituenti maggiori relativamente flessibile, variabile o libero con possibilità di scrambling pragmaticamente motivato, topicalizzazione sintattica, focalizzazione contrastiva o altre permutazioni dell'ordine basico canonico per scopi discorsivi, informazionali o enfatici, tendono sistematicamente e robustatamente a mantenere, preservare e a volte anche elaborare ulteriormente sistemi di marcatura casuale morfologica più robusti, ricchi, articolati ed estesi sia sui pronomi sia sui nomi comuni, perché la marcatura morfologica esplicita, visibile e univoca diventa funzionalmente necessaria, comunicativamente indispensabile e cognitivamente vantaggiosa per disambiguare e identificare chiaramente le funzioni sintattiche e i ruoli semantici quando l'ordine lineare superficiale delle parole non è rigido, fisso, predicibile e affidabile come unico indicatore delle relazioni grammaticali.

Formalmente, possiamo esprimere questo universale come probabilità condizionale quantitativa: P(VO | Caso_solo_pronominale) ≈ 0,90-0,95, ovvero la probabilità che una lingua sia tipologicamente VO dato che ha marcatura casuale morfologica limitata e ristretta al sistema pronominale è molto alta, dell'ordine del 90-95% secondo stime empiriche. Equivalentemente formulato nella direzione inversa dell'implicazione: P(Caso_solo_pronominale | OV) ≈ 0,05-0,10, ovvero la probabilità che una lingua abbia marcatura del caso limitata solo ai pronomi dato che è tipologicamente OV è molto bassa, solo 5-10%, indicando che la stragrande maggioranza delle lingue OV hanno sistemi casuali più ricchi estesi anche ai nomi comuni.

L'italiano moderno standard contemporaneo fornisce un esempio particolarmente chiaro, trasparente e paradigmatico di questo pattern tipologico correlazionale. L'italiano è una lingua inequivocabilmente e prototipicamente SVO (quindi sottotipo di VO) con ordine basico relativamente fisso e rigido dei costituenti maggiori verbali e nominali, permettendo solo variazioni marginali, stilisticamente marcate o pragmaticamente condizionate dall'ordine canonico SVO non marcato. La marcatura morfologica esplicita del caso nominale è estremamente limitata, drasticamente ridotta e virtualmente atrofizzata nell'italiano contemporaneo rispetto al latino classico da cui l'italiano deriva diacronicamente attraverso evoluzioni fonetiche, morfologiche e sintattiche complesse e multistrato.

La distinzione nominativo-accusativo, che rappresenta la distinzione casuale morfologica più fondamentale, primaria e universalmente attestata nelle lingue con allineamento sintattico nominativo-accusativo (distinguendo il soggetto come argomento unico del verbo intransitivo e argomento agentivo del verbo transitivo dall'oggetto come argomento paziente del verbo transitivo), è visibile morfologicamente ed esplicitamente differenziata solo in alcuni pronomi personali tonici forti e in alcuni clitici pronominali deboli: "egli" (nominativo soggetto maschile singolare terza persona, forma tonica forte arcaizzante) / "lui" (forma obliqua usata modernamente anche come nominativo colloquiale) / "lo" (clitico accusativo maschile

singolare); "ella" (nominativo femminile singolare terza persona, arcaizzante) / "lei" (obliquo/nominativo moderno) / "la" (clitico accusativo femminile); "io" (nominativo prima persona singolare) / "me" (accusativo/dativo/obliquo forma tonica) / "mi" (clitico accusativo/dativo); "tu" (nominativo seconda persona singolare) / "te" (accusativo/dativo/obliquo) / "ti" (clitico); "noi" (invariabile per nominativo/accusativo/obliquo) / "ci" (clitico); "voi" (invariabile) / "vi" (clitico).

Per i nomi comuni lessicali pieni, al contrario assoluto e categorico, non esiste alcuna marcatura morfologica flessiva distinguibile, visibile o produttiva che differenzi morfologicamente il caso nominativo soggetto dal caso accusativo oggetto: le frasi "Marco vede Lucia" (Marco soggetto agente, Lucia oggetto paziente) e "Lucia vede Marco" (Lucia soggetto agente, Marco oggetto paziente) hanno significati radicalmente, completamente e diametralmente diversi (attori invertiti, relazioni semantiche invertite) derivanti totalmente, esclusivamente e univocamente dall'ordine lineare sequenziale delle parole secondo la grammatica configurazionale rigida SVO dell'italiano, con il primo sintagma nominale nella sequenza lineare interpretato grammaticalmente, categoricamente e senza ambiguità come soggetto agente e il secondo sintagma nominale interpretato come oggetto paziente secondo la struttura configurazionale rigida e predicibile.

Questa dipendenza totale, assoluta ed esclusiva dall'ordine fisso delle parole per indicare, codificare e disambiguare le funzioni sintattiche fondamentali è resa possibile funzionalmente, comunicativamente adeguata e cognitivamente non problematica dal fatto cruciale che l'italiano possiede e mantiene un ordine basico canonico SVO relativamente rigido, fisso e predicibile che non permette liberamente scrambling arbitrario degli argomenti maggiori, permutazioni casuali o inversioni pragmatiche senza condizioni morfosintattiche e discorsive molto specifiche, restrittive e marcate.

Questo pattern strutturale e tipologico si allinea perfettamente, conferma precisamente e esemplifica paradigmaticamente la previsione derivante dall'Universale 41 di Greenberg: l'italiano ha marcatura casuale morfologica drasticamente limitata e ristretta al sistema pronominale (distinguendo nominativo/accusativo solo nei pronomi ma non nei nomi comuni) ed è effettivamente e inequivocabilmente una lingua VO, specificamente SVO, con ordine relativamente rigido che rende la marcatura casuale morfologica estesa funzionalmente meno necessaria o addirittura ridondante.

Il francese moderno standard mostra un pattern ancora più estremo, radicale e avanzato di atrofia casuale morfologica progressiva: anche i pronomi soggetto personali tonici sono morfologicamente quasi completamente invariabili, indistinguibili o non differenziati foneticamente dai loro corrispondenti accusativi/obliqui in molte varietà parlate colloquiali contemporanee ("je" nominativo versus "me" accusativo sono fonologicamente distinti [ʒə] versus [mə], ma "il" nominativo versus "le" accusativo sono identici o indistinguibili foneticamente in pronuncia standard [il] versus [lə] che si riducono spesso a [l] proclitico identico), rendendo l'ordine rigido fisso delle parole ancora più cruciale, indispensabile e funzionalmente centrale per l'interpretazione sintattica e semantica corretta delle relazioni grammaticali.

Al contrario diretto, speculare e complementare, lingue prototipicamente e armonicamente SOV come il turco ottomano e moderno, il giapponese classico e contemporaneo, il

coreano, l'hindi-urdu, il tamil e altre lingue dravidiche meridionali hanno tipicamente, sistematicamente e robustatamente sistemi di marcatura casuale morfologica molto più sviluppati, ricchi, articolati, elaborati ed estesi morfologicamente non solo ai pronomi ma anche e soprattutto ai nomi comuni lessicali pieni attraverso affissi flessivi produttivi, sistematici e morfologicamente trasparenti.

Il turco moderno standard, lingua altaica meridionale prototipicamente SOV e armonicamente head-final, marca morfologicamente ed esplicitamente l'oggetto diretto definito e specifico mediante il suffisso casuale accusativo -ı/-i/-u/-ü (con varianti allomorfiche condizionate dalle regole sistematiche, predicibili e produttive di armonia vocalica palatale/velare anteriore/posteriore e arrotondata/non-arrotondata che governano pervasivamente la fonologia morfofonologica turca): "Kitabı gördüm" (libro-ACCUSATIVO vedere-PASSATO-1SG = "ho visto il libro" / "vidi il libro"), dove la radice nominale "kitab" (libro) riceve obbligatoriamente il suffisso flessivo accusativo -ı per indicare esplicitamente, morfologicamente e univocamente che è l'oggetto diretto definito del verbo transitivo "gördüm" (vidi).

Il soggetto nominativo, al contrario e complementarmente, appare tipicamente e prototipicamente senza marcatura flessiva casuale esplicita morfologicamente visibile quando è al caso nominativo non marcato o caso zero (forma basica non suffissata): "Kitap burada" (libro qui = "il libro è qui"), "Adam geldi" (uomo venire-PASSATO = "l'uomo venne"), dove i sostantivi "kitap" (libro) e "adam" (uomo) appaiono nella loro forma basica non flessa senza suffissi casuali espliciti, interpretati automaticamente come soggetti nominativi per default.

Tuttavia, e questo è il punto cruciale teoricamente e tipologicamente significativo per la validità dell'Universale 41, il turco marca morfologicamente ed esplicitamente anche molti altri casi grammaticali e semantici oltre al semplice nominativo e accusativo, inclusi: genitivo -ın/-in/-un/-ün (possessore, relazione parte-tutto), dativo -a/-e (meta, recipiente, beneficiario), locativo -da/-de (locazione statica, tempo), ablativo -dan/-den (origine, provenienza, agente passivo, comparazione), strumentale-comitativo -la/-le (strumento, compagnia), creando un sistema ricco, articolato e produttivo di sei casi morfologicamente distinti, differenziati e sistematicamente marcati più varie postposizioni specializzate che governano e richiedono casi specifici secondo reggenze idiomatiche.

Esempi sistematici attraverso il paradigma casuale completo: "kitap" (nominativo, soggetto/forma basica), "kitab-ı" (accusativo, oggetto diretto definito), "kitab-ın" (genitivo, del libro), "kitab-a" (dativo, al libro), "kitab-da" (locativo, nel/sul libro), "kitab-dan" (ablativo, dal libro), "kitab-la" (strumentale, con il libro). Questo paradigma completo, ricco e sistematico si applica produttivamente e regolarmente a virtualmente tutti i nomi comuni lessicali, non solo ai pronomi, fornendo marcatura morfologica esplicita estensiva.

Questo sistema elaborato, ricco e morfologicamente esplicito di marcatura casuale permette e facilita funzionalmente al turco una certa flessibilità significativa, pragmaticamente motivata e discorsivamente condizionata nell'ordine lineare superficiale delle parole per scopi pragmatici, informativi o enfatici: l'ordine basico canonico non marcato è rigorosamente SOV, ma permutazioni come OSV, topicalizzazioni contrastive, focalizzazioni enfatiche e scrambling parziale sono grammaticalmente possibili, accettabili e

comunicativamente efficaci perché la funzione sintattica è codificata morfologicamente ed esplicitamente attraverso i suffissi casuali oltre che posizionalmente attraverso l'ordine lineare.

Per esempio, "Kitabı gördüm ben" (libro-ACCUSATIVO vedere-PASSATO-1SG io = "il libro, l'ho visto IO") permuta e estrae l'oggetto accusativo "kitabı" in posizione iniziale topicalizzata per enfasi contrastiva o topicalizzazione discorsiva, ma la marcatura accusativa esplicita -ı garantisce morfologicamente, univocamente e senza ambiguità che "kitabı" sia interpretato correttamente come oggetto diretto paziente nonostante la posizione lineare superficiale non canonica e inusuale. Senza la marcatura accusativa morfologica, questa permutazione sarebbe potenzialmente ambigua o mal interpretabile.

Similmente: "Ben kitabı gördüm" (io libro-ACCUSATIVO vedere-PASSATO-1SG, ordine SOV canonico non marcato), "Kitabı ben gördüm" (libro-ACCUSATIVO io vedere-PASSATO-1SG, oggetto topicalizzato enfaticamente), "Gördüm ben kitabı" (vedere-PASSATO-1SG io libro-ACCUSATIVO, ordine fortemente marcato con verbo iniziale per enfasi massima), tutte queste permutazioni sono grammaticalmente legittime, interpretabili univocamente e comunicativamente efficaci grazie alla marcatura casuale morfologica esplicita che disambigua le funzioni sintattiche indipendentemente dall'ordine posizionale.

Il giapponese classico e moderno manifesta un pattern strutturalmente simile, tipologicamente parallelo e funzionalmente convergente ma con un sistema morfologico diverso basato su particelle postposizionali invariabili piuttosto che suffissi flessivi agglutinanti. Le funzioni sintattiche fondamentali sono indicate sistematicamente ed esplicitamente da particelle grammaticali invariabili che seguono immediatamente i sintagmi nominali: "が" (ga) marca il soggetto nominativo/caso grammaticale principale; "を" (o/wo) marca l'oggetto diretto accusativo; "に" (ni) marca il dativo-locativo (recipiente, meta, locazione); "で" (de) marca lo strumentale-locativo (strumento, mezzo, locazione dell'azione); "の" (no) marca il genitivo-possessivo; "と" (to) marca il comitativo-congiuntivo (compagnia, congiunzione); "から" (kara) marca l'ablativo (origine, provenienza); "まで" (made) marca il terminativo (limite finale spaziale/temporale).

Esempio paradigmatico: "太郎が本を読んだ" (Tarō ga hon o yonda = Taro NOMINATIVO libro ACCUSATIVO leggere-PASSATO = "Taro lesse il/un libro"), dove le particelle "ga" e "o" marcano esplicitamente e univocamente le funzioni sintattiche di soggetto e oggetto rispettivamente, permettendo interpretazione non ambigua.

Questa marcatura esplicita sistematica mediante particelle permette al giapponese scrambling relativamente libero e flessibilità nell'ordine dei costituenti per scopi pragmatici, stilistici o enfatici: "本を太郎が読んだ" (hon o Tarō ga yonda = libro ACCUSATIVO Taro NOMINATIVO leggere-PASSATO) permuta l'oggetto in posizione iniziale topicalizzata ma mantiene interpretazione univoca grazie alle particelle casuali; "読んだ太郎が本を" (yonda Tarō ga hon o = leggere-PASSATO Taro NOMINATIVO libro ACCUSATIVO, ordine fortemente marcato e stilisticamente particolare ma grammaticalmente possibile).

L'ordine basico preferito e statisticamente dominante rimane SOV secondo studi quantitativi su corpora naturali, ma deviazioni motivate pragmaticamente sono grammaticalmente lecite, comunicativamente efficaci e cognitivamente processabili proprio grazie alla marcatura

casuale esplicita morfologica che disambigua le relazioni sintattiche indipendentemente dalla posizione lineare.

Il rapporto sistematico, robusto e cross-linguisticamente attestato tra genitivi e adposizioni, già ampiamente discusso, analizzato e documentato in relazione all'Universale 2, rivela ulteriori pattern correlazionali complessi e implicazionali teoricamente significativi che arricchiscono, approfondiscono e articolano ulteriormente il quadro teorico complessivo delle correlazioni tipologiche. Greenberg osserva con precisione empirica, accuratezza descrittiva e perspicacia teorica che nelle lingue con preposizioni dominanti (adposizioni che precedono sistematicamente il loro complemento nominale secondo ordine Preposizione-Nome), il modificatore genitivo possessivo tende robustatamente, sistematicamente e predicibilmente a seguire il nome governante o testa nominale posseduta, generando costruzioni morfosintattiche trasparenti e analiticamente chiare come "il libro di Marco" in italiano, "the book of John" in inglese, "le livre de Jean" in francese, "el libro de Juan" in spagnolo, "o livro de João" in portoghese, dove la preposizione genitiva specializzata ("di" / "of" / "de") introduce esplicitamente il modificatore genitivo possessore e lo marca sintatticamente come dipendente grammaticale del nome testa precedente.

Queste costruzioni genitive preposizionali risultano particolarmente trasparenti, analiticamente chiare e facilmente analizzabili dal punto di vista sintattico compositivo perché l'elemento adposizionale funzionale (la preposizione genitiva "di"/"of"/"de") marca esplicitamente, morfologicamente e univocamente la relazione di dipendenza sintattica, possessione semantica o associazione concettuale tra il nome possessore/modificatore (Marco, John, Jean, Juan, João) e il nome posseduto/testa (libro, book, livre), fungendo da connettore funzionale esplicito e morfologicamente indipendente.

Questa marcatura esplicita attraverso preposizioni specializzate, funzionalmente dedicate e grammaticalizzate contrasta nettamente, strutturalmente e tipologicamente con le costruzioni genitive caratteristiche delle lingue con postposizioni dominanti o sistemi morfologici casuali ricchi, dove la relazione genitiva possessiva è tipicamente, prevalentemente e sistematicamente marcata attraverso morfologia flessiva diretta applicata al nome dipendente stesso (suffissi o desinenze casuali genitive che si agglutinano o si fondono con la radice nominale), e il genitivo morfologicamente marcato precede rigidamente e categoricamente la testa nominale posseduta senza necessaria mediazione o interposizione di elementi adposizionali indipendenti e fonologicamente separati.

Nel turco, come già osservato, analizzato e documentato precedentemente in relazione all'Universale 2, il genitivo possessivo è espresso morfologicamente attraverso il suffisso casuale genitivo -ın/-in/-un/-ün che si agglutina direttamente alla radice nominale secondo le regole sistematiche di armonia vocalica: "Mehmet'in kitabı" (Mehmet-GENITIVO libro-POSSESSIVO.3SG = "il libro di Mehmet"), dove il genitivo morfologicamente marcato mediante suffissazione "Mehmet'in" precede rigidamente, categoricamente e invariabilmente il nome testa "kitap", in perfetta armonia strutturale e direzionale con il pattern generale OV della lingua (l'oggetto precede il verbo nella frase transitiva basica) e con la presenza sistematica, dominante e produttiva di postposizioni piuttosto che preposizioni come categoria adposizionale caratteristica.

Notare inoltre che il nome posseduto "kitap" riceve simultaneamente un suffisso possessivo concordante -ı (terza persona singolare) che marca morfologicamente la relazione di possessione dal lato del nome posseduto, creando una doppia marcatura morfologica esplicita: il possessore riceve marcatura genitiva, il posseduto riceve marcatura possessiva concordante, garantendo chiarezza referenziale massima.

Gerarchie di Animacy, Definiteness e Marcatura Differenziale: Pattern Universali Complessi

Un dominio empiricamente ricco, tipologicamente variabile, teoricamente profondo e universalmente significativo riguarda l'interazione complessa e multidimensionale tra proprietà semantiche referenziali dei sintagmi nominali (animacy/inanimacy, definiteness/indefiniteness, specificità/non-specificità, topicality/focus) e la marcatura morfologica differenziale di ruoli sintattici e semantici, fenomeno conosciuto tecnicamente come Differential Object Marking (DOM) o più generalmente Differential Argument Marking (DAM).

La gerarchia di animacy o scala di animatezza rappresenta una scala universale o quasi-universale di prominenza cognitiva, saliency comunicativa e topicality discorsiva che ordina gerarchicamente le entità nominali secondo il loro grado di animatezza, agency potenziale e prominenza cognitiva per i parlanti umani. Una formulazione standard ampiamente accettata della gerarchia di animacy è:

1ª persona > 2ª persona > 3ª persona umana > Animali > Oggetti inanimati concreti > Astratti

oppure in forma più articolata e dettagliata:

Parlante (1ª pers) > Ascoltatore (2ª pers) > Umani noti/propri > Umani comuni > Animali superiori > Animali inferiori > Piante > Oggetti naturali inanimati > Oggetti artificiali > Concetti astratti

Questa gerarchia non è arbitraria, convenzionale o culturalmente specifica ma riflette plausibilmente principi cognitivi universali profondi della cognizione umana, dell'attenzione percettiva e dell'organizzazione concettuale: gli esseri umani, e particolarmente i partecipanti diretti dell'atto comunicativo (parlante e ascoltatore), sono cognitivamente più salienti, comunicativamente più topicali e referencialmente più centrali rispetto ad entità non umane, inanimate o astratte.

Questa gerarchia manifesta effetti sistematici, robusti e cross-linguisticamente attestati in numerosi fenomeni grammaticali indipendenti: marcatura casuale differenziale, accordo verbale, accessibilità alla relativizzazione, gerarchie di controllo, binding anaforico, marcatura di voce (attivo/passivo), patterns di incorporazione nominale, ordine preferenziale dei costituenti, e molti altri domini grammaticali.

La marcatura differenziale dell'oggetto (Differential Object Marking, DOM) è un fenomeno tipologicamente diffuso, cross-linguisticamente robusto e teoricamente significativo dove la marcatura morfologica esplicita dell'oggetto diretto (mediante caso accusativo, preposizioni

specializzate, o altre strategie morfosintattiche) dipende sistematicamente da proprietà semantiche e pragmatiche del sintagma nominale oggetto piuttosto che essere applicata uniformemente, categoricamente e indiscriminatamente a tutti gli oggetti diretti indipendentemente dalle loro proprietà referenziali.

Tipicamente e cross-linguisticamente, oggetti diretti alti nella gerarchia di animacy (particolarmente oggetti umani, definiti, specifici, topicali, pronominali) ricevono marcatura morfologica esplicita obbligatoria o preferita, mentre oggetti bassi nella gerarchia (inanimati, indefiniti, non-specifici, non-topicali, lessicalmente pieni) ricevono marcatura zero, ridotta o opzionale.

Lo spagnolo standard contemporaneo fornisce un esempio paradigmatico, chiaro e ben studiato di DOM. Lo spagnolo marca morfologicamente gli oggetti diretti umani, animati e definiti mediante la preposizione specializzata accusativa "a" (la stessa forma della preposizione dativa-locativa ma funzionalmente distinta come marcatore accusativo), chiamata tecnicamente "a personal" o "a accusativa":

"Veo a Juan" (vedere.1SG A Juan = "vedo Juan"), con marcatura esplicita obbligatoria mediante "a" per oggetto umano proprio definito.

"Veo a la profesora" (vedo A la professoressa = "vedo la professoressa"), con "a" obbligatoria per oggetto umano definito comune.

Ma: "Veo el libro" (vedo il libro = "vedo il libro"), *"Veo a el libro" è agrammaticale, senza "a" per oggetto inanimato.

"Veo un perro" (vedo un cane = "vedo un cane"), senza "a" per animale indefinito, ma "Veo a mi perro" (vedo A mio cane) con "a" per animale posseduto definito topicale.

Il turco mostra DOM sistematica e morfologicamente trasparente attraverso il caso accusativo: oggetti diretti definiti e specifici ricevono obbligatoriamente il suffisso accusativo -ı/-i/-u/-ü, mentre oggetti indefiniti e non-specifici appaiono nella forma basica non marcata (caso nominativo o assoluto):

"Kitab-ı okudum" (libro-ACCUSATIVO leggere-PASSATO-1SG = "lessi IL libro", definito specifico).

"Kitap okudum" (libro∅ leggere-PASSATO-1SG = "lessi UN libro / lessi libri", indefinito non-specifico).

Il persiano moderno mostra DOM attraverso la particella postposizionale "rā" che marca obbligatoriamente oggetti diretti definiti, specifici e topicali ma è omessa per oggetti indefiniti:

"ketāb rā xāndam" (libro RA leggere-PASSATO-1SG = "lessi IL libro").

"ketāb xāndam" (libro∅ leggere-PASSATO-1SG = "lessi UN libro").

L'hindi-urdu manifesta DOM estremamente elaborata e condizionata da multiple proprietà: la postposizione accusativa "ko" è obbligatoria per oggetti umani, altamente preferita per animati definiti, opzionale per inanimati definiti, e assente per indefiniti non-specifici:

"mãi laṛke ko dekhtā hũ̃" (io ragazzo KO vedere.ABITUALE essere.1SG = "io vedo IL ragazzo"), obbligatoria per umano.

"mãi kutte ko dekhtā hũ̃" (io cane KO vedere = "vedo IL cane"), preferita per animale definito.

"mãi kitāb (ko) dekhtā hũ̃" (io libro (KO) vedere = "vedo IL libro"), opzionale per inanimato.

"mãi kitāb dekhtā hũ̃" (io libro∅ vedere = "vedo UN libro"), assente per indefinito.

Questi pattern di DOM possono essere teoricamente modellati attraverso scale implicazionali gerarchiche che predicono la distribuzione della marcatura:

Se oggetti inanimati ricevono marcatura esplicita, allora oggetti animati ricevono marcatura (ma non viceversa).

Se oggetti indefiniti ricevono marcatura, allora oggetti definiti ricevono marcatura (ma non viceversa).

Se nomi comuni ricevono marcatura, allora pronomi ricevono marcatura (ma non viceversa).

Formalmente: Inanimato → Animato → Umano → Pronominale, dove → indica implicazione della marcatura lungo la gerarchia.

Spiegazioni Teoriche Integrate: Verso una Sintesi Multidimensionale

Gli universali linguistici identificati da Greenberg e confermati, raffinati, estesi e quantificati dalla ricerca tipologica successiva sollevano questioni teoriche fondamentali, profonde e centrali che stanno al cuore della scienza del linguaggio e della cognizione umana: quali sono i fattori causali, motivazionali o esplicativi che determinano, producono o favoriscono questi pattern cross-linguistici non casuali, sistematici e robusti? Perché certi ordini, combinazioni e configurazioni di proprietà grammaticali sono sistematicamente favoriti, altamente frequenti e cross-linguisticamente attestati mentre altri sono rari, marginali o completamente assenti? Cosa ci dicono queste regolarità empiriche impressionanti sulla natura intrinseca della facoltà linguistica umana, sui principi organizzativi che governano la struttura del linguaggio e sui vincoli che delimitano lo spazio delle grammatiche possibili?

Nella letteratura linguistica teorica contemporanea, sono state proposte, elaborate, difese e dibattute fondamentalmente quattro categorie principali di spiegazioni teoriche, non necessariamente mutualmente esclusive ma potenzialmente complementari, sinergiche e integrabili in modelli teorici più ricchi e multidimensionali:

1. Spiegazioni Innatiste: Grammatica Universale, Vincoli Geneticamente Determinati e Principi Innati

La prospettiva innatista, associata principalmente e storicamente con la tradizione generativista chomskyana iniziata negli anni 1950-1960 e sviluppata attraverso molteplici

incarnazioni teoriche successive (teoria standard, teoria standard estesa, governo e legamento, principi e parametri, minimalismo), propone che gli universali linguistici riflettano fondamentalmente vincoli intrinseci, biologicamente determinati e geneticamente codificati sulla facoltà linguistica umana (Language Faculty), una componente specializzata, modulare e specie-specifica della cognizione umana codificata nel programma genetico dello sviluppo neurologico specifico della specie Homo sapiens.

Secondo questa visione teorica influente e ampiamente dibattuta, la Grammatica Universale (Universal Grammar, UG) costituisce un sistema innato, geneticamente determinato e biologicamente specificato di principi universali invarianti e parametri binari o discreti di variazione che definisce, delimita e restringe drasticamente lo spazio delle grammatiche naturali umane possibili, escludendo a priori configurazioni logicamente concepibili ma biologicamente impossibili.

I principi universali sono affermazioni assolute, inviolabili e senza eccezioni che si applicano obbligatoriamente, necessariamente e universalmente a tutte le lingue naturali umane senza eccezioni, derivando dalla natura intrinseca della facoltà linguistica geneticamente specificata. Esempi proposti di principi universali includono: Principio di Proiezione (argomenti tematici richiesti semanticamente devono essere proiettati sintatticamente), Principio di Categorie Vuote (tracce di movimento devono essere propriamente governate), Condizione di Soggiacenza (movimento non può attraversare più di una barriera), Condizione C (espressioni referenziali non possono essere c-comandate e legate da antecedenti), tra molti altri principi tecnici proposti in diverse fasi della teoria generativa.

I parametri rappresentano punti dimensionali di variazione binaria, discreta o finitamente multivalente dove le lingue possono differire in modi limitati, finiti e strutturati secondo opzioni prestabilite biologicamente. L'esempio paradigmatico classico è il parametro pro-drop o parametro del soggetto nullo: alcune lingue permettono omissione del soggetto pronominale quando recuperabile dal contesto e dall'accordo verbale ricco (italiano, spagnolo, greco, giapponese, cinese), mentre altre richiedono soggetti sintatticamente espressi obbligatoriamente (inglese, francese, tedesco standard).

Il parametro testa-complemento (head-directionality parameter) specificherebbe se in una data lingua le teste sintattiche precedono o seguono sistematicamente i loro complementi, con questa scelta binaria applicata coerentemente attraverso categorie sintattiche diverse (V-O versus O-V, P-N versus N-P, N-Gen versus Gen-N). Fissare questo parametro su un valore particolare durante l'acquisizione linguistica avrebbe automaticamente conseguenze in cascata per molteplici aspetti dell'ordine delle parole, generando le correlazioni osservate empiricamente come universali implicazionali di Greenberg.

La forza predittiva e il potere esplicativo dell'approccio parametrico risiedono nel fatto teoricamente attraente che un numero relativamente piccolo di parametri binari o discreti (proposte classiche suggerivano 20-50 parametri fondamentali) può generare combinatorialmente una vasta tipologia di lingue attraverso le loro combinazioni possibili (2^n configurazioni per n parametri binari indipendenti), mentre simultaneamente escludendo sistematicamente, categoricamente e a priori certe combinazioni impossibili, proibite o fortemente dispreferite che violano vincoli derivanti dalle interazioni parametriche o dai principi universali.

Critiche e problemi con l'approccio puramente innatista includono: primo, il carattere statistico, graduale e probabilistico piuttosto che assoluto e deterministico della maggior parte degli universali di Greenberg, con eccezioni sistematiche e frequenze intermedie che sono difficili da riconciliare con vincoli genetici assoluti (se gli universali derivassero direttamente da vincoli geneticamente codificati rigidi, ci aspetteremmo universali assoluti al 100% senza eccezioni); secondo, la difficoltà di formulare parametri che catturano precisamente e completamente le correlazioni tipologiche osservate senza overgenerazione o undergenerazione; terzo, problemi di acquisizione parametrica (bootstrapping problem, trigger problem); quarto, la natura altamente astratta e non direttamente osservabile dei principi e parametri proposti, rendendo difficile la verifica empirica diretta.

2. Spiegazioni Funzionaliste: Pressioni Comunicative, Principi di Efficienza e Adattamento Funzionale

L'approccio funzionalista, sviluppato particolarmente nelle tradizioni tipologiche europee, cognitive e usage-based, propone che gli universali linguistici riflettano fondamentalmente adattamenti a pressioni funzionali derivanti dall'uso del linguaggio come strumento di comunicazione sociale, trasmissione di informazione e interazione coordinata. Le lingue, secondo questa visione, evolvono culturalmente e si strutturano storicamente in modi che ottimizzano l'efficienza comunicativa, minimizzano l'ambiguità referenziale, facilitano la produzione e la comprensione in tempo reale, e massimizzano la trasmissione accurata dell'informazione nelle condizioni ecologiche reali, variabili e complesse dell'interazione comunicativa umana.

Diversi principi funzionali specifici sono stati proposti, formulati e testati empiricamente per spiegare particolari universali implicazionali:

Principio di Dipendenza Localizzata (Locality of Dependencies): Elementi sintatticamente o semanticamente correlati, interdipendenti o mutuamente selezionati tendono ad essere collocati linearmente vicini nella sequenza superficiale per facilitare il processamento incrementale, minimizzare il carico sulla memoria di lavoro limitata e ridurre l'ambiguità temporanea durante il parsing. Questo principio predice che complementi strettamente selezionati semanticamente dalle loro teste (come oggetti diretti argomenti selezionati dal verbo transitivo) appariranno adiacenti o vicini alle teste selezionanti, mentre modificatori più loosely associati, opzionali o aggiunti possono apparire più distanti.

Principio dell'Ordine Iconico (Iconicity Principle): L'ordine lineare sequenziale delle espressioni linguistiche nella serializzazione superficiale tende a riflettere iconicamente, isomorficamente e trasparentemente l'ordine temporale, causale, logico o concettuale degli eventi, stati o relazioni semantiche descritte. Per esempio, nella sequenza prototipica Soggetto-Verbo-Oggetto (SVO), l'agente causativo (soggetto) che causa, inizia e controlla l'azione precede linearmente l'azione stessa (verbo) che precede l'entità affetta, modificata o creata (oggetto), riflettendo iconicamente la catena causale sequenziale tipica degli eventi transitivi prototipici.

Principio della Minimizzazione Computazionale (Minimize Computational Load): Configurazioni grammaticali che minimizzano la complessità computazionale necessaria per la produzione incrementale e il parsing online sono sistematicamente preferite, favorite

evolutivamente e culturalmente trasmesse con maggiore successo rispetto a configurazioni che richiedono rianalisi sintattica, backtracking computazionale o mantenimento prolungato di dipendenze non risolte attraverso lunghe distanze lineari con materiale interveniente.

Principio della Marcatura Informazionale (Information Structure Principle): Informazione più topicale, data, presupposta o accessibile contestualmente tende ad apparire prima, inizialmente o prominentemente nella frase, mentre informazione più focale, nuova, assertita o communicativamente centrale appare dopo, finalmente o in posizioni strutturalmente prominenti. Questo principio può motivare l'ordine preferenziale universale Soggetto > Oggetto, dato che il soggetto è prototipicamente e statisticamente più topicale, più dato e più accessibile dell'oggetto in contesti discorsivi non marcati secondo studi quantitativi su corpora cross-linguistici.

L'armonia direzionale (head-directionality harmony), da prospettiva funzionalista, può essere spiegata come conseguenza naturale della minimizzazione del carico computazionale: essere consistenti nella direzione di ramificazione attraverso diverse categorie sintattiche riduce la complessità procedurale per il parser incrementale, che può applicare la stessa strategia di attaccamento strutturale, le stesse euristiche di predizione e le stesse routines computazionali uniformemente attraverso categorie piuttosto che dover monitorare, tracciare e applicare strategie diverse, context-sensitive e categoria-specifiche.

3. Spiegazioni Psicolinguistiche: Vincoli del Processamento, Architetture Cognitive e Limitazioni Computazionali

La prospettiva psicolinguistica, strettamente correlata ma distinguibile dall'approccio funzionalista, enfatizza i vincoli specifici, limitazioni quantitative e biases sistematici imposti dai meccanismi cognitivi concreti e architetture neurali specifiche del processamento linguistico in tempo reale: percezione fonetica, accesso lessicale, parsing sintattico incrementale, produzione articolatoria, memoria di lavoro limitata, attenzione selettiva. Questi vincoli derivano dalle architetture computazionali specifiche, limitazioni di capacità quantitative e caratteristiche di processamento dei sistemi cognitivi umani che implementano neuralmente la facoltà linguistica.

Effetti di Peso Sintattico e Ramificazione Pesante (Heavy-NP Shift, End-Weight): Costituenti sintatticamente pesanti, complessi e articolati (lunghi in numero di parole, profondamente annidati in struttura sintattica, morfologicamente complessi) tendono ad essere collocati verso la fine della frase o del sintagma per facilitare la pianificazione incrementale nella produzione, permettere buffering temporale durante formulazione, e ridurre il carico sulla memoria di lavoro limitata durante il parsing comprehensivo incrementale. Questo principio spiega perché clausole relative, essendo costituenti intrinsecamente pesanti, complessi e potenzialmente molto lunghi, tendono a seguire pattern di peso: lingue right-branching le collocano a destra (postnominali) in armonia con ramificazione generale, lingue left-branching possono collocarle a sinistra ma con maggiore tensione cognitiva e limitazioni di lunghezza.

Preferenza per Dipendenze Brevi (Dependency Locality Theory, DLT): Il parser umano preferisce minimizzare la distanza lineare (misurata in parole intervenenti) e strutturale (misurata in nodi sintattici intervenenti) tra teste sintattiche e dipendenti grammaticalmente

correlati. Dipendenze lunghe impongono maggior carico sulla memoria di lavoro, aumentano la probabilità di errori di parsing e garden-path effects, e rallentano la comprensione. Questo principio può motivare ordini armonici dove teste e complementi sono adiacenti o molto vicini linearly.

Effetti di Sorpresa Informazionale e Aspettativa Predittiva (Surprisal, Expectation, Entropy Reduction): Il processamento è facilitato, accelerato e reso più accurato quando le strutture sintattiche incontrate corrispondono alle aspettative generate predittivamente dal contesto precedente attraverso modelli probabilistici impliciti acquisiti dall'esperienza linguistica. Ordini consistenti, predicibili e statisticamente frequenti riducono la sorpresa informazionale (misurata in teoria dell'informazione come log-probabilità inversa) e facilitano il processamento predittivo. Armonia direzionale può emergere come pattern ottimale che massimizza la predicibilità strutturale cross-categoriale.

Evidenza empirica sistematica da esperimenti psicolinguistici controllati supporta parzialmente queste spiegazioni processuali. Studi di eye-tracking durante la lettura mostrano che violazioni di pattern armonici causano rallentamenti misurabili nel tempo di lettura, aumento delle regressioni oculari (movimenti saccadici all'indietro) e maggior tempo totale di processamento, suggerendo maggior difficoltà cognitiva. Studi di potenziali evento-correlati (ERP, Event-Related Potentials) mostrano componenti elettrofisiologiche di violazione sintattica (P600, LAN) quando ordini non armonici o inaspettati sono presentati a parlanti nativi.

Tuttavia, queste evidenze sono metodologicamente confuse dal fatto critico che parlanti sperimentali sono nativi di lingue particolari specifiche con patterns armonici specifici, e quindi le loro risposte di processamento riflettono inevitabilmente sia vincoli cognitivi universali sia l'influenza dell'esperienza specifica con il sistema grammaticale acquisito nativamente. Distinguere vincoli universali di processamento da effetti di abitudine linguistica language-specific richiede studi cross-linguistici attentamente controllati su parlanti di lingue tipologicamente diverse, che rimangono relativamente scarsi, metodologicamente difficili e teoricamente ambigui.

4. Spiegazioni Diacroniche: Percorsi di Cambiamento, Deriva Tipologica e Evoluzione Culturale

La prospettiva diacronica ed evolutiva propone che gli universali sincronici osservati nella distribuzione tipologica contemporanea delle lingue riflettano vincoli sistematici, biases direzionali e canalizzazioni evolutive sui processi di cambiamento linguistico storico, trasmissione culturale intergenerazionale e deriva tipologica. Certi percorsi evolutivi, transizioni tipologiche e traiettorie di cambiamento sono intrinsecamente più probabili, naturali e frequentemente attestati di altri a causa di meccanismi di rianalisi, grammaticalizzazione unidirezionale, analogia e livellamento che operano sistematicamente attraverso generazioni di acquisizione imperfetta e trasmissione culturale modificativa.

Grammaticalizzazione Unidirezionale: Percorsi di grammaticalizzazione (processo per cui elementi lessicali sostantivi/verbi diventano progressivamente elementi grammaticali funzionali, e elementi grammaticali diventano più grammaticali, più affissali o più ridotti fonologicamente) seguono traiettorie sistematiche, predicibili e cross-linguisticamente

ricorrenti. Per esempio, preposizioni derivano storicamente e diacronicamente da verbi seriali, sostantivi relazionali o costruzioni possessive secondo percorsi ben documentati, e questa origine etimologica può favorire, predire o vincolare certi ordini posizionali piuttosto che altri.

Deriva Tipologica Coerente e Cambiamento Sistemico (Typological Drift): Lingue in evoluzione diacronica tendono a muoversi progressivamente verso configurazioni tipologiche coerenti, armoniche e internamente consistenti piuttosto che verso configurazioni miste, disarmoniche o internamente contraddittorie. Una lingua che comincia ad acquisire alcune proprietà VO attraverso cambiamenti iniziali isolati (per esempio, riordinamento verbo-oggetto da OV originario a VO innovativo attraverso processi sintattici graduali) tenderà progressivamente, sistematicamente e su scale temporali lunghe ad acquisire anche altre proprietà VO correlate e armoniche (cambio da postposizioni originarie a preposizioni innovative, da genitivi prenominali a postnominali, da clausole relative prenominali a postnominali) attraverso analogia proporzionale, livellamento paradigmatico e deriva sistemica, generando drift tipologica coerente verso un tipo più armonicamente consistente.

Canalizzazione Evolutiva e Trasmissione Iterata (Evolutionary Channeling, Iterated Learning): La trasmissione linguistica attraverso generazioni successive, mediata dall'acquisizione imperfetta da parte dei bambini con input limitato e degenerato, può funzionare come filtro selettivo che sistematicamente favorisce configurazioni più regolari, trasparenti, composizionali e facili da acquisire. Pattern irregolari, opachi, disarmonici o idiosincratici possono essere progressivamente regolarizzati, sistematizzati o eliminati attraverso generazioni successive di acquisizione imperfetta e rianalisi creativa, producendo deriva verso regolarità, armonia e sistematicità.

Evidenza per spiegazioni diacroniche proviene da studi dettagliati di cambiamento linguistico storicamente documentato. L'inglese ha cambiato drasticamente da ordine più OV, flessivamente ricco e morfologicamente sintetico nell'inglese antico (Old English, circa 450-1150 d.C.) a ordine più rigidamente VO, morfologicamente impoverito e analiticamente trasparente nell'inglese moderno (Modern English, dal 1500 circa), e questo cambiamento sistemico è stato accompagnato da cambiamenti correlati armonici: perdita progressiva della flessione casuale nominale, perdita della flessione casuale aggettivale, ordine progressivamente più rigido SVO, sviluppo di costruzioni progressive e passive più analitiche.

Studi di simulazione computazionale usando modelli multi-agente di popolazioni linguistiche in evoluzione culturale mostrano che armonia tipologica può emergere spontaneamente e auto-organizzarsi anche senza vincoli innati espliciti codificati, semplicemente attraverso dinamiche di trasmissione culturale iterata e apprendimento con piccole preferenze o biases di apprendibilità che favoriscono marginalmente pattern armonici.

Conclusioni: Integrazione Teorica, Prospettive Future e Significato Profondo degli Universali

La ricerca sugli universali linguistici iniziata pionieristicamente da Joseph Greenberg nel suo lavoro seminale del 1963 e continuata intensamente, estesa quantitativamente e approfondita teoricamente nei decenni successivi attraverso il lavoro di generazioni di tipologi, sintatticisti formali, psicolinguisti e linguisti storici ha rivelato pattern cross-linguistici impressionanti, robusti e sistematici che richiedono spiegazione teorica articolata, multidimensionale e empiricamente fondata.

La distribuzione non casuale, altamente asimmetrica e statisticamente significativa delle proprietà grammaticali nella tipologia mondiale, le correlazioni implicazionali forti, robuste e predicibili tra domini sintattici apparentemente distinti e indipendenti, e le lacune tipologiche sistematiche, persistenti e non accidentali nello spazio delle grammatiche logicamente possibili ma empiricamente non attestate costituiscono fatti empirici fondamentali, robusti e teoricamente profondi che qualsiasi teoria adeguata, empiricamente responsabile e scientificamente seria del linguaggio umano deve affrontare, spiegare e derivare teoricamente.

È progressivamente chiaro, emergente dal consenso crescente nella comunità scientifica e supportato da evidenza empirica convergente da molteplici domini indipendenti, che una spiegazione completamente adeguata, empiricamente soddisfacente e teoricamente profonda degli universali linguistici richiederà probabilmente, necessariamente e inevitabilmente un'integrazione sintetica, multidimensionale e teoricamente articolata di molteplici prospettive teoriche complementari piuttosto che l'adozione esclusiva, riduzionista e monolitica di un singolo framework esplicativo privilegiato.

Vincoli innati sulla facoltà linguistica derivanti dalla specificità genetica specie-specifica, pressioni comunicative funzionali derivanti dall'uso sociale del linguaggio, limitazioni architetturali del processamento cognitivo in tempo reale derivanti dalle capacità computazionali finite dei sistemi neurali, e dinamiche evolutive del cambiamento diacronico derivanti dalla trasmissione culturale intergenerazionale probabilmente tutti contribuiscono in modi interagenti, sinergici e mutuamente rafforzanti alla determinazione, modellazione e mantenimento dei pattern osservati empiricamente nella tipologia cross-linguistica mondiale.

Un modello teorico integrato e multidimensionale potrebbe concettualizzare gli universali linguistici come emergenti dall'interazione complessa di molteplici livelli causali:

Livello 1 - Vincoli Biologici Innati: Principi universali assoluti derivanti dalla struttura geneticamente determinata della facoltà linguistica (per esempio, dipendenza strutturale piuttosto che lineare nelle trasformazioni sintattiche, principio di proiezione, principi di legamento), che definiscono rigidamente lo spazio delle grammatiche biologicamente possibili escludendo a priori configurazioni incompatibili con l'architettura computazionale innata.

Livello 2 - Biases di Processamento Cognitivo: Preferenze soft derivanti dalle architetture neurali del processamento (preferenza per dipendenze locali, peso finale, riduzione dell'entropia predittiva), che non escludono rigidamente configurazioni ma le rendono più difficili, costose o instabili, creando pressioni selettive graduali.

Livello 3 - Pressioni Funzionali Comunicative: Ottimizzazione emergente verso efficienza comunicativa, riduzione dell'ambiguità, trasmissione accurata dell'informazione, che favorisce evolutivamente lingue con certe proprietà strutturali.

Livello 4 - Dinamiche Diacroniche e Trasmissione Culturale: Canalizzazione evolutiva attraverso percorsi di grammaticalizzazione, deriva tipologica coerente, trasmissione iterata con bias di apprendibilità, che produce convergenza verso attrattori tipologici.

Questi livelli non operano indipendentemente ma interagiscono sinergicamente: vincoli innati delimitano lo spazio possibile, biases di processamento creano gradienti di difficoltà all'interno di questo spazio, pressioni funzionali favoriscono regioni ottimali, e dinamiche diacroniche producono deriva verso queste regioni attraverso tempo evolutivo.

Direzioni Future di Ricerca: Sfide Metodologiche e Opportunità Empiriche

Un programma di ricerca futuro produttivo, empiricamente rigoroso e teoricamente progressivo sugli universali linguistici dovrebbe perseguire multiple direzioni complementari e sinergiche:

1. Espansione della Copertura Empirica Cross-Linguistica

Raccogliere dati sistematici, metodologicamente controllati e comparabilmente codificati su campioni sempre più ampi, geneticamente diversificati e geograficamente rappresentativi di lingue, includendo particolarmente lingue poco documentate, endangered, non scritte o parlate in comunità piccole e remote che potrebbero rivelare pattern tipologici precedentemente sconosciuti, eccezioni significative agli universali proposti o configurazioni intermedie teoricamente informative.

Il World Atlas of Language Structures (WALS), sebbene rappresenti un progresso monumentale con oltre 2600 lingue e 200 feature tipologiche sistematicamente codificate, copre ancora solo circa il 40% delle circa 7000 lingue attualmente parlate nel mondo. Molte lingue indigenous, minoritarie o endangered rimangono drasticamente sottodocumentate o completamente non descritte tipologicamente. L'espansione della documentazione linguistica sistematica rappresenta una priorità urgente sia per preservazione culturale sia per completezza scientifica della tipologia.

Database tipologici digitali, freely accessible, standardizzati metodologicamente e continuamente aggiornati come WALS Online, Grambank, AUTOTYP, e altri progetti collaborativi internazionali rappresentano infrastrutture cruciali per ricerca quantitativa rigorosa. L'integrazione di questi database attraverso ontologie condivise, coding schemes compatibili e API interoperabili permetterebbe meta-analisi potenti e test statistici più robusti.

2. Quantificazione Precisa delle Correlazioni e Modellazione Statistica Sofisticata

Andare oltre affermazioni qualitative, impressionistiche o binarie di correlazioni implicazionali verso misurazioni quantitative precise, statisticamente rigorose e probabilisticamente informative di forze associative usando strumenti statistici sofisticati, modelli grafici probabilistici, analisi di reti tipologiche e machine learning, permettendo discriminazione

fine-grained tra universali forti, moderati e deboli, identificazione di fattori modulatori e confounding variables, e quantificazione dell'incertezza nelle stime.

Tecniche statistiche avanzate come regressione logistica multivariata, modelli gerarchici bayesiani con controllo filogenetico per dipendenza genetica, analisi di reti complesse, modelli di equazioni strutturali e causal inference methods possono rivelare pattern più sottili, interazioni complesse tra variabili, effetti mediatori e moderatori, e relazioni causali direzionali piuttosto che semplici correlazioni.

Controllo per dipendenza filogenetica (genetic relatedness bias) e areale (contact-induced similarity) attraverso metodi filogenetici comparativi adattati dalla biologia evolutiva (phylogenetic comparative methods, PCMs) permette di distinguere tendenze universali genuine da pattern dovuti a eredità comune o diffusione areale.

3. Testing Sperimentale di Predizioni Teoriche attraverso Metodologie Multiple

Derivare predizioni testabili specifiche, falsificabili e quantitative da diverse teorie esplicative concorrenti e progettare studi empirici rigorosi, metodologicamente controllati e replicabili (esperimenti psicolinguistici cross-linguistici, analisi di corpora naturali multi-lingue, simulazioni computazionali agent-based, studi di acquisizione infantile cross-linguistica, studi neurolinguistici comparativi) che permettano discriminazione empirica tra ipotesi alternative, competitive e teoricamente distinte.

Psicolinguistica Cross-Linguistica Comparativa: Condurre esperimenti di processamento (self-paced reading, eye-tracking, ERP, fMRI) su parlanti nativi di lingue tipologicamente diverse testate con materiali comparabili tradotti per valutare se violazioni di armonia tipologica producono difficoltà di processamento universali o language-specific. Se difficoltà emergono universalmente anche in parlanti di lingue che permettono gli ordini "violativi", questo supporta vincoli universali di processamento. Se difficoltà emergono solo in parlanti di lingue armoniche, questo supporta effetti di esperienza specifica.

Studi di Acquisizione: Tracciare ordini temporali di acquisizione di diverse costruzioni in bambini acquisitori di lingue tipologicamente diverse. Se predizioni derivanti da gerarchie implicazionali (elemento meno marcato acquisito prima di elemento più marcato) sono confermate cross-linguisticamente, questo supporta la validità cognitiva delle gerarchie.

Simulazioni Computazionali: Modelli agent-based di popolazioni linguistiche in evoluzione culturale con apprendimento iterato (Iterated Learning Models, ILMs) possono testare se pattern tipologici osservati emergono spontaneamente da dinamiche di trasmissione culturale con bias minimi di apprendibilità, oppure richiedono vincoli innati specifici pre-cablati.

4. Modellazione Formale Computazionalmente Esplicita e Teoricamente Articolata

Sviluppare modelli formali matematicamente rigorosi, computazionalmente espliciti e implementabili che integrano molteplici fattori (vincoli strutturali, pressioni funzionali, dinamiche di apprendimento, evoluzione culturale) e generano predizioni quantitative precise sulla distribuzione tipologica attesa, frequenze relative di diversi tipi, probabilità di transizioni diacroniche e patterns di variazione sincronica.

Modelli computazionali basati su teoria dell'informazione, ottimizzazione Bayesiana, dynamical systems, evolutionary game theory, e statistical learning possono formalizzare rigorosamente intuizioni teoriche verbali e testarle quantitativamente contro dati empirici.

Per esempio, modelli di comunicazione efficiente basati su rate-distortion theory predicono trade-offs ottimali tra complessità espressiva e effort articolatorio, che possono derivare predizioni quantitative su lunghezza parole, distribuzione Zipfian, marcatura differenziale.

5. Investigazione Sistematica della Variazione: Eccezioni, Casi Marginali e Configurazioni Intermedie

Studiare non solo le tendenze centrali, modali o prototipiche ma anche e soprattutto le eccezioni, i casi marginali, le lingue tipologicamente miste, le configurazioni intermedie e le transizioni in corso, che possono rivelare limitazioni importanti delle generalizzazioni, fattori modulatori non precedentemente considerati, condizioni di applicabilità ristrette e meccanismi causali più sottili.

Eccezioni sistematiche e regolari sono spesso più informative teoricamente di conformità ovvie. Il cinese mandarino come "eccezione" all'Universale 24 ha generato decenni di dibattito teorico produttivo che ha raffinato sostanzialmente la nostra comprensione della natura delle clausole relative, della distinzione tra modificatori nominali diversi e della complessità della tipologia sintattica.

Lingue tipologicamente miste come persiano, amarico, varie lingue ethiopiche, alcune lingue caucasiche forniscono laboratori naturali per studiare come lingue reali gestiscono tensioni tra principi organizzativi concorrenti e raggiungono compromessi strutturali.

6. Integrazione Multi-Dimensionale: Connettere Tipologia, Acquisizione, Processamento, Cambiamento e Neurobiologia

Connettere sistematicamente ricerca tipologica con ricerca in acquisizione linguistica infantile (come bambini acquisiscono costruzioni tipologicamente diverse? le gerarchie implicazionali predicono ordini di acquisizione?), psicolinguistica sperimentale (pattern tipologici correlano con facilità di processamento?), linguistica storica e cambiamento linguistico (transizioni tipologiche seguono percorsi predicibili? esistono attrattori tipologici stabili?), e neurobiologia del linguaggio (basi neurali diverse per lingue tipologicamente diverse? plasticità neurale nell'acquisizione di L2 tipologicamente distante?).

Questa integrazione multi-livello può rivelare se universali operano coerentemente attraverso diversi livelli di analisi (acquisizione, processamento, cambiamento, implementazione neurale) suggerendo principi profondamente radicati, oppure se diversi pattern emergono a livelli diversi suggerendo meccanismi causali multipli indipendenti.

Significato Filosofico e Implicazioni per la Scienza Cognitiva

Gli universali linguistici, lungi dall'essere semplici curiosità empiriche, regolarità descrittive o generalizzazioni tassonomiche, costituiscono una finestra privilegiata, metodologicamente unica e teoricamente profonda sulla natura intrinseca della facoltà linguistica umana, rivelando i principi organizzativi fondamentali, i vincoli strutturali innati o emergenti, e le

pressioni selettive multiple che modellano l'evoluzione, l'acquisizione, l'uso e la struttura di questo sistema cognitivo straordinariamente complesso, sofisticato, onnipresente e centrale all'esperienza umana.

La capacità linguistica rappresenta una caratteristica definitoria, specie-specifica e culturalmente centrale di Homo sapiens, distinguendoci qualitativamente da tutte le altre specie animali nonostante sistemi comunicativi animali sofisticati (canti degli uccelli, calls dei primati, danze delle api). La ricchezza strutturale illimitata, la creatività generativa infinita, la dipendenza dalla struttura piuttosto che dalla sequenza, la ricorsività illimitata e la trasmissione culturale cumulativa del linguaggio umano non hanno paralleli conosciuti nel regno animale.

Gli universali linguistici ci informano su aspetti invarianti, necessari o altamente probabili di questa capacità distintiva. Essi delimitano lo spazio delle possibilità, rivelano tendenze sistematiche e identificano principi organizzativi che potrebbero riflettere:

- Proprietà interne della architettura computazionale geneticamente specificata della facoltà linguistica (prospettiva innatista/generativista)

- Adattamenti funzionali a esigenze comunicative universali derivanti dalla natura sociale, cooperativa e comunicativa dell'esistenza umana (prospettiva funzionalista/comunicativa)

- Vincoli derivanti dalle limitazioni computazionali dei sistemi neurali finiti che implementano il processamento linguistico (prospettiva psicolinguistica/procedurale)

- Attrattori emergenti nella evoluzione culturale dei sistemi linguistici attraverso trasmissione generazionale iterata (prospettiva evolutiva/diacronica)

La questione cruciale non è quale singola prospettiva sia "corretta" in senso esclusivo, ma piuttosto come queste prospettive complementari possono essere integrate sinteticamente in un framework unificato che riconosce molteplici livelli causali operanti simultaneamente e interattivamente.

Conclusione Finale: L'Eredità di Greenberg e la Ricerca Continua

L'eredità duratura, profonda e transformativa di Joseph Harold Greenberg risiede non solo negli specifici universali implicazionali concreti che identificò pionieristicamente nel suo lavoro seminale del 1963, molti dei quali rimangono empiricamente robusti, statisticamente significativi e teoricamente influenti dopo sei decenni di ricerca intensiva, scrutinio critico e testing empirico estensivo su campioni molto più ampi, ma anche e soprattutto nel metodo empirico rigoroso, comparativo sistematico, quantitativamente orientato e cross-linguisticamente informato che introdusse, perfezionò e esemplificò nella linguistica tipologica.

Questo metodo ha trasformato radicalmente, permanentemente e irrevocabilmente lo studio della variazione linguistica da un'impresa largamente descrittiva, tassonomica, idiografica e particolaristica focalizzata su descrizioni dettagliate di lingue individuali in una scienza

nomotetica, generalizzante, teoreticamente ambiziosa e empiricamente rigorosa che cerca leggi, principi e patterns generali sottostanti alla diversità superficiale apparentemente caotica.

Greenberg dimostrò convincentemente che è possibile e produttivo cercare regolarità significative, testare ipotesi quantitative e formulare generalizzazioni cross-linguisticamente valide nonostante la stupefacente diversità superficiale delle lingue umane. Egli mostrò che tipologia sistematica informata teoricamente e teoria formale empiricamente fondata non sono approcci antagonistici, incompatibili o mutuamente esclusivi ma possono e devono essere integrati sinergicamente.

La continua esplorazione degli universali linguistici attraverso metodologie sempre più sofisticate, rigorose e multidisciplinari, dataset sempre più comprensivi, rappresentativi e accurati, e frameworks teorici sempre più articolati, formalmente espliciti e empiricamente testabili promette di approfondire significativamente, sostanzialmente e progressivamente la nostra comprensione scientifica di cosa significa essere creature dotate del dono straordinario, misterioso e definitorio del linguaggio umano.

Gli universali linguistici continuano a rappresentare un dominio di ricerca vivace, dinamico, metodologicamente innovativo e teoricamente fertile che connette intimamente linguistica teorica, tipologia empirica, psicolinguistica sperimentale, acquisizione infantile, cambiamento diacronico, neurobiologia cognitiva e filosofia del linguaggio in un'impresa scientifica integrata, multidisciplinare e profondamente illuminante sulla natura umana.

Bibliografia Essenziale Selezionata e Annotata

Opere Fondative di Greenberg:

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of Language* (pp. 73-113). MIT Press. *L'articolo seminale originale che ha definito il campo moderno degli universali linguistici e introdotto i 45 universali implicazionali basati su analisi di 30 lingue.*

Greenberg, J. H. (1966). *Language Universals, with Special Reference to Feature Hierarchies*. Mouton. *Estensione e raffinamento teorico dell'approccio universalista con enfasi su gerarchie implicazionali.*

Trattati Tipologici Classici:

Comrie, B. (1981). *Language Universals and Linguistic Typology: Syntax and Morphology*. University of Chicago Press. *Trattazione sistematica, pedagogicamente eccellente e empiricamente ricca della tipologia sintattica e morfologica con focus su universali implicazionali.*

Croft, W. (2003). *Typology and Universals* (2nd ed.). Cambridge University Press. *Sintesi teorica moderna che integra prospettive funzionaliste, tipologiche e formali con discussione approfondita di metodologia e spiegazioni.*

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, 68(1), 81-138. *Analisi quantitativa rigorosa e teoricamente sofisticata delle correlazioni di ordine delle parole su campioni ampi con test statistici dettagliati.*

Hawkins, J. A. (1983). *Word Order Universals*. Academic Press. *Trattazione completa delle correlazioni di ordine con proposta teorica basata su principi di processamento e peso sintattico.*

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford University Press. *Teoria generale degli universali basata su principi di efficienza computazionale e complessità nel processamento.*

Atlanti Tipologici e Database:

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (Eds.). (2005). *The World Atlas of Language Structures*. Oxford University Press. *Collezione monumentale di 142 mappe tipologiche basate su 2560 lingue, risorsa empirica fondamentale.*

Dryer, M. S., & Haspelmath, M. (Eds.). (2013). *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. http://wals.info/ *Versione online continuamente aggiornata con nuove features e lingue.*

Prospettive Teoriche Formali:

Baker, M. C. (2001). *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Basic Books. *Presentazione accessibile della teoria parametrica generativista applicata alla variazione tipologica.*

Newmeyer, F. J. (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford University Press. *Tentativo sistematico di riconciliare tipologia empirica e teoria generativa formale.*

Cinque, G. (Ed.). (2002). *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Volume 1*. Oxford University Press. *Approccio cartografico alla struttura funzionale con implicazioni per universali di ordine.*

Approcci Funzionalisti e Cognitivi:

Givón, T. (1979). *On Understanding Grammar*. Academic Press. *Articolazione influente dell'approccio funzionalista agli universali grammaticali.*

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. University of Chicago Press. *Studio comparativo massiccio di grammaticalizzazione con implicazioni per universali diacronici.*

Croft, W., & Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge University Press. *Introduzione alla linguistica cognitiva con applicazioni tipologiche.*

Psicolinguistica e Processamento:

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95-126). MIT Press. *Teoria formale del processamento sintattico con predizioni per universali basati su complessità computazionale.*

Hawkins, J. A. (2014). *Cross-Linguistic Variation and Efficiency*. Oxford University Press. *Sintesi matura della teoria dell'efficienza nel processamento applicata a universali tipologici.*

Evoluzione e Cambiamento:

Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge University Press. *Teoria usage-based dell'emergenza di struttura grammaticale e cambiamento.*

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681-10686. *Evidenza sperimentale per emergenza di struttura linguistica attraverso trasmissione iterata.*

Marcatura Differenziale:

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435-483. *Analisi formale influente di DOM in prospettiva Optimality Theory.*

Bossong, G. (1985). *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Gunter Narr Verlag. *Studio pioneristico sistematico di marcatura differenziale dell'oggetto.*

Metodologia Quantitativa:

Bickel, B. (2007). Typology in the 21st century: Major current developments. *Linguistic Typology*, 11(1), 239-251. *Riflessione metodologica sullo stato e futuro della tipologia quantitativa.*

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473(7345), 79-82. *Applicazione di metodi filogenetici comparativi a universali di ordine delle parole con controllo per dipendenza genetica.*